OM protein - protein search, using sw model

Run on: March 20, 2000, 05:31:20 ; Search time 35.25 Seconds
(without alignments)
249.964 Million cell updates/sec

Title: US-08-509-359B-138
Perfect score: 1923
Sequence: 1 EELTLKYGAKHVIMLFVPVT.....STDNLVRPFMDTLASHQLYI 372

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A_Geneseq_36:*

Word size : 0

Number of hits that pass the threshold : 188963

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1923	100.0	448	1	W05762	Human presenilin-2
2	1923	100.0	448	1	W11321	Human AD4 protein.
3	1923	100.0	448	1	W23967	Human presenilin-2
4	1919	99.8	448	1	W05763	Presenilin-2 M239V
5	1918	99.7	448	1	W05765	Presenilin-2 I420T
6	1914	99.5	448	1	W05764	Presenilin-2 N141I
7	1907.5	99.2	447	1	W28508	Full AD4/AD3LP seq
8	1723	89.6	414	1	W05766	Presenilin-2 delta
9	1591.5	82.8	376	1	W28506	AD4/AD3LP sequence
10	1431.5	74.4	467	1	W05735	Murine presenilin.
11	1431.5	74.4	467	1	W23966	Mouse presenilin-1
12	1429.5	74.3	467	1	W11839	Human early onset
13	1429.5	74.3	463	1	W11840	Early onset Alzhei
14	1429.5	74.3	467	1	W05733	Presenilin-1-1. Ne
15	1429.5	74.3	407	1	W28507	Partial AD3 sequen
16	1429.5	74.3	467	1	W41430	PS1/467 protein. D
17	1429.5	74.3	429	1	W41429	PS1/429 protein. D
18	1429.5	74.3	467	1	W23964	Human presenilin-1
19	1426.5	74.2	467	1	W05755	Presenilin-1-1 L28
20	1426.5	74.2	467	1	W05758	Presenilin-1-1 L39
21	1426.5	74.2	467	1	W05737	Presenilin-1-1 V82
22	1426.5	74.2	467	1	W05746	Presenilin-1-1 I21
23	1426.5	74.2	467	1	W41431	Mouse PS1/467 prot
24	1425.5	74.1	467	1	W05754	Presenilin-1-1 A28
25	1425.5	74.1	467	1	W05736	Presenilin-1-1 A79
26	1425.5	74.1	467	1	W05747	Presenilin-1-1 I23
27	1425.5	74.1	467	1	W05749	Presenilin-1-1 A26
28	1424.5	74.1	463	1	W12376	Human S182 gene pr
29	1424.5	74.1	467	1	W05738	Presenilin-1-1 V96
30	1424.5	74.1	467	1	W05739	Presenilin-1-1 Y11
31	1424.5	74.1	467	1	W05741	Presenilin-1-1 I14
32	1424.5	74.1	467	1	W05748	Presenilin-1-1 A24
33	1424.5	74.1	467	1	W27176	Human S182 gene, P
34	1423.5	74.0	467	1	W05753	Presenilin-1-1 E28
35	1423.5	74.0	467	1	W05757	Presenilin-1-1 G38
36	1423.5	74.0	463	1	W05734	Presenilin-1-2. Ne
37	1423.5	74.0	467	1	W05740	Presenilin-1-1 M13
38	1423.5	74.0	467	1	W05742	Presenilin-1-1 M14
39	1423.5	74.0	463	1	W23965	Human presenilin-1
40	1422.5	74.0	467	1	W05744	Presenilin-1-1 L17
41	1422.5	74.0	467	1	W56770	Homo sapiens PS-1.
42	1421.5	73.9	467	1	W05752	Presenilin-1-1 P26
43	1421.5	73.9	463	1	W42375	Human presenilin I
44	1420.5	73.9	467	1	W05743	Presenilin-1-1 H16
45	1420.5	73.9	467	1	W05745	Presenilin-1-1 G20

ALIGNMENTS

RESULT 1
W05762
ID W05762 standard; Protein; 448 AA.
AC W05762;
DT 25-JUL-1997 (first entry)
DE Human presenilin-2.
KW Presenilin-2; human; hPS1-1; hPS1-2; PS-2; integral membrane protein; AD;
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW depression; antibody; gene expression modulator; therapy.
OS Homo sapiens.
PN WO9634099-A2.
PD 31-OCT-1996.
PF 29-APR-1996; CA0263.
PR 28-APR-1995; US-431048.
PR 28-JUN-1995; US-496841.
PR 31-JUL-1995; US-509359.
PA (HSCR-) HSC RES & DEV LP.
PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
PI Fraser PE, Rommens JM, St George-Hyslop PH;
DR WPI; 96-497631/49.
DR N-PSDB; T40031.
PT New presenilin genes - useful for diagnosis, therapy and drug
PT screening of familial Alzheimer's disease, cerebral disorders, etc.
PS Claim 4; Page 148-150; 178pp; English.
CC This sequence represents the wild type human presenilin-2.
CC W05733 and W05734 represent the two different forms of wild type human
CC presenilin-1 (PS-1). The form represented by W05734 results from
CC alternate splicing of the genomic DNA sequence. W05735 represents the
CC coding sequence for wild type mouse PS-1. The presenilins are a family of
CC highly conserved integral membrane proteins with a common structural
CC motif, common alternate splicing patterns, and common mutational hot spot
CC regions. Mutations in PS genes are implicated in familial Alzheimer's
CC disease (AD) and possibly other diseases such as cerebral haemorrhage,
CC schizophrenia, depression etc., so detection of mutations in the DNA
CC encoding these sequences can be used for diagnosis of these diseases.
CC These proteins, or vectors that express them or containing antisense
CC sequences, antibodies selective for mutant forms of these proteins (such
CC as W05736) and modulators of PS gene expression are potentially useful
CC for treatment of AD etc. Transgenic animals are useful as models for drug
CC screening. The antibodies can also be used e.g. for affinity purification
CC and in immunoassays.
SQ Sequence 448 AA;

Query Match 100.0%; Score 1923; DB 1; Length 448;
Best Local Similarity 100.0%; Pred. No. 1.6e-200;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 136

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 196

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     197 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 256

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
FT   Domain          225. .244
FT                   /label= TM5
FT                   /note= "transmembrane domain 5"
FT   Domain          245. .249
FT                   /label= TM5-6
FT                   /note= "hydrophilic loop"
FT   Domain          250. .268
FT                   /label= TM6
FT                   /note= "transmembrane domain 6"
FT   Domain          269. .387
FT                   /label= TM6-7
FT                   /note= "hydrophilic loop"
FT   Domain          388. .409
FT                   /label= TM8
FT                   /note= "transmembrane domain 8"
FT   Misc_difference 141
FT                   /note= "Asn141Ile mutation site (Claim 19)"
FT   Misc_difference 239
FT                   /note= "Met239Val mutation site (Claim 19)"
FT   Misc_difference 420
FT                   /note= "Ile420Thr mutation site"
PN   WO9801549-A2.
PD   15-JAN-1998.
PF   04-JUL-1997; CA0475.
PR   02-JAN-1997; US-034590.
PR   05-JUL-1996; US-021673.
PR   12-JUL-1996; US-021700.
PR   08-NOV-1996; US-029895.
PA   (HSCR-) HSC RES & DEV LP.
PA   (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
PI   Fraser PE, Rommens JM, St George-Hyslop PH;
DR   WPI; 98-286355/25.
DR   N-PSDB; V04669.
PT   New isolated mutant presenilin-1 genes - useful for developing
PT   products for use in detection, diagnosis and therapy of Alzheimer's
PT   disease and for drug screening
PS   Claim 19; Page 203-204; 238pp; English.
CC   This polypeptide comprises human presenilin-2 (hPS2).  Its amino
CC   acid sequence was deduced from an isolated cDNA clone (see V04669).
CC   Human and murine presenilin-1 sequences are also provided (see
CC   W23964-66).  Mutations in the PS-1 and PS-2 genes are linked to
CC   the development in humans of forms of familial Alzheimer's disease
CC   (FAD) and may be causative of other disorders, e.g. cognitive,
CC   intellectual, neurological or physiological disorders such as
CC   cerebral haemorrhage, schizophrenia, depression, mental retardation
CC   and epilepsy.  Use of the nucleic acids and proteins comprising or
CC   derived from the presenilins is made in screening and diagnosing
CC   FAD, identifying and developing therapeutics for treatment of FAD,
CC   and in producing cell lines and transgenic animals useful as models
CC   of FAD.  Methods for identifying substances that bind to, or
CC   modulate the activity of a presenilin protein, and methods for
CC   identifying substances that affect the interaction of a
CC   presenilin-interacting protein with a presenilin protein are also
CC   disclosed.
SQ   Sequence   448 AA;

  Query Match             100.0%;  Score 1923;  DB 1;  Length 448;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-200;
  Matches  372;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 136

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 196

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     197 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 256

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     257 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 316

Qy     241 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     317 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 376

Qy     301 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     377 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 436

Qy     361 FMDTLASHQLYI 372
           ||||||||||||
Db     437 FMDTLASHQLYI 448


RESULT   4
W05763
ID   W05763 standard; Protein; 448 AA.
AC   W05763;
DT   25-JUL-1997  (first entry)
DE   Presenilin-2 M239V mutation.
KW   Presenilin-2; human; hPS1-1; hPS1-2; PS-2; integral membrane protein; AD;
KW   familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW   depression; antibody; gene expression modulator; therapy; mutein.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   modified_site   239
FT                   /label= M239V
PN   WO9634099-A2.
PD   31-OCT-1996.
PF   29-APR-1996; CA0263.
PR   28-APR-1995; US-431048.
PR   28-JUN-1995; US-496841.
PR   31-JUL-1995; US-509359.
PA   (HSCR-) HSC RES & DEV LP.
PA   (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
PI   Fraser PE,  Rommens JM,  St George-Hyslop PH;
DR   WPI; 96-497631/49.
PT   New presenilin genes - useful for diagnosis, therapy and drug
PT   screening of familial Alzheimer's disease, cerebral disorders, etc.
PS   Claim 4; Page -; 178pp; English.
CC   W05763-W05766 represent mutated versions of the human presenilin-2
CC   protein (see W05762 for wild type sequence). The presenilins are a family
CC   of highly conserved integral membrane proteins with a common structural
CC   motif, common alternate splicing patterns, and common mutational hot spot
CC   regions. Mutations in PS genes are implicated in familial Alzheimer's
CC   disease (AD) and possibly other diseases such as cerebral haemorrhage,
CC   schizophrenia, depression etc., so detection of mutations in the DNA
CC   encoding the wild type sequences can be used for diagnosis of these
CC   diseases. The wild type proteins, or vectors that express them or
CC   containing antisense sequences, antibodies selective for these mutant
CC   forms of the proteins and modulators of PS gene expression are
CC   potentially useful for treatment of AD etc. Transgenic animals are useful
CC   as models for drug screening. The antibodies can also be used e.g. for
CC   affinity purification and in immunoassays.
SQ   Sequence   448 AA;

  Query Match             99.8%;  Score 1919;  DB 1;  Length 448;
  Best Local Similarity   99.7%;  Pred. No. 4.3e-200;
  Matches  371;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 136

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 196

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
```

```
Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137 NSVLITLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 196

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     197 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 256

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     257 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 316

Qy     241 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     317 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 376

Qy     301 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     377 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 436

Qy     361 FMDTLASHQLYI 372
           ||||||||||||
Db     437 FMDTLASHQLYI 448
```

RESULT 7
W28508
ID W28508 standard; Protein; 447 AA.
AC W28508;
DT 07-DEC-1997 (first entry)
DE Full AD4/AD3LP sequence.
KW AD3; AD4/AD3LP; Alzheimer's disease; chromosome; missegregation;
KW presenilin; inhibitor; AD; trisomy 21; ss.
OS Homo sapiens.
PN WO9707213-A2.
PD 27-FEB-1997.
PF 15-AUG-1996; U13314.
PR 16-AUG-1995; US-002448.
PA (HARD) HARVARD COLLEGE.
PI Li J, Potter H;
DR WPI; 97-165297/15.
DR N-PSDB; T87426.
PT Identifying genes which cause chromosome missegregation - useful for
PT identifying causes of and treatments for diseases, e.g. Alzheimer's
PT disease, cancer and ageing
PS Claim 29; Fig 29; 77pp; English.
CC Identifying genes which cause improper chromosome segregation,
CC screening for inhibitors of chromosome missegregation and processes
CC caused by genes encoding chromosome missegregation promoters
CC was exemplified using Alzheimer's disease. The sequences
CC given in T87401 to T87426 can be used in the above methods.
CC It is not clear from the figure legend, the figure and the
CC disclosure of the specification which sequence of Fig 1 and Fig 28
CC is the AD4/AD3LP or the AD3 sequence.
SQ Sequence 447 AA;

Query Match 99.2%; Score 1907.5; DB 1; Length 447;
Best Local Similarity 99.7%; Pred. No. 7.6e-199;
Matches 371; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 136

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 196

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     197 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 256

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     257 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 316

Qy     241 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 300
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     317 LPYDPEM-EDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 375

Qy     301 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     376 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 435

Qy     361 FMDTLASHQLYI 372
           ||||||||||||
Db     436 FMDTLASHQLYI 447
```

RESULT 8
W05766
ID W05766 standard; Protein; 414 AA.
AC W05766;
DT 25-JUL-1997 (first entry)
DE Presenilin-2 delta263-296 mutation.
KW Presenilin-2; human; hPS1-1; hPS1-2; PS-2; integral membrane protein; AD;
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW depression; antibody; gene expression modulator; therapy; mutein.
OS Homo sapiens.
FH Key Location/Qualifiers
FT misc_difference 263. .264
FT /note= "site of 34 residue deletion"
PN WO9634099-A2.
PD 31-OCT-1996.
PF 29-APR-1996; CA0263.
PR 28-APR-1995; US-431048.
PR 28-JUN-1995; US-496841.
PR 31-JUL-1995; US-509359.
PA (HSCR-) HSC RES & DEV LP.
PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
PI Fraser PE, Rommens JM, St George-Hyslop PH;
DR WPI; 96-497631/49.
PT New presenilin genes - useful for diagnosis, therapy and drug
PT screening of familial Alzheimer's disease, cerebral disorders, etc.
PS Claim 4; Page -; 178pp; English.
CC W05763-W05766 represent mutated versions of the human presenilin-2
CC protein (see W05762 for wild type sequence). The presenilins are a family
CC of highly conserved integral membrane proteins with a common structural
CC motif, common alternate splicing patterns, and common mutational hot spot
CC regions. Mutations in PS genes are implicated in familial Alzheimer's
CC disease (AD) and possibly other diseases such as cerebral haemorrhage,
CC schizophrenia, depression etc., so detection of mutations in the DNA
CC encoding the wild type sequences can be used for diagnosis of these
CC diseases. The wild type proteins, or vectors that express them or
CC containing antisense sequences, antibodies selective for these mutant
CC forms of the proteins and modulators of PS gene expression are
CC potentially useful for treatment of AD etc. Transgenic animals are useful
CC as models for drug screening. The antibodies can also be used e.g. for
CC affinity purification and in immunoassays.
SQ Sequence 414 AA;

Query Match 89.6%; Score 1723; DB 1; Length 414;
Best Local Similarity 90.9%; Pred. No. 7.4e-179;
Matches 338; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

```
Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 136

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 196
```

```
Qy    241 LPYDPEME--------EDSYDSFGEPSYPEVFEPPLTGYPG-----------EEL---- 277
          :| :|:          :||     :  : | :|     : |            :||
Db    309 VPKNPKYNTQRAERETQDSGSGNDDGGFSEEWEAQRDSHLGPHRSTPESRAAVQELSGSI 368

Qy    277 ---EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVF 333
             |: ||||||||||||||||||||:|| |||||||:||||||||||||||||||:|
Db    369 LTSEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIF 428

Qy    334 KKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
          ||||||||||||||:|||:|| ||:|||| || || ||
Db    429 KKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467
```

RESULT 11
W23966
ID W23966 standard; Protein; 467 AA.
AC W23966;
DT 20-JUL-1998 (first entry)
DE Mouse presenilin-1 homologue.
KW Presenilin-1; PS1 gene; mouse; familial Alzheimer's disease; FAD;
KW cerebral haemorrhage; schizophrenia; depression; epilepsy;
KW mental retardation; diagnosis; therapy; transgenic animal.
OS Mus musculus.

FH	Key	Location/Qualifiers
FT	Domain	82. .100
FT		/label= TM1
FT		/note= "transmembrane domain 1"
FT	Domain	101. .132
FT		/label= TM1-2
FT		/note= "hydrophilic loop"
FT	Domain	133. .154
FT		/label= TM2
FT		/note= "transmembrane domain 2"
FT	Domain	155. .163
FT		/label= TM2-3
FT		/note= "hydrophilic loop"
FT	Domain	164. .183
FT		/label= TM3
FT		/note= "transmembrane domain 3"
FT	Domain	184. .194
FT		/label= TM3-4
FT		/note= "hydrophilic loop"
FT	Domain	195. .212
FT		/label= TM4
FT		/note= "transmembrane domain 4"
FT	Domain	213. .220
FT		/label= TM4-5
FT		/note= "hydrophilic loop"
FT	Domain	221. .238
FT		/label= TM5
FT		/note= "transmembrane domain 5"
FT	Domain	239. .243
FT		/label= TM5-6
FT		/note= "hydrophilic loop"
FT	Domain	244. .262
FT		/label= TM6
FT		/note= "transmembrane domain 6"
FT	Domain	263. .407
FT		/label= TM6-7
FT		/note= "hydrophilic loop"
FT	Domain	408. .428
FT		/label= TM8
FT		/note= "transmembrane domain 8"
FT	Misc_difference	177
FT		/note= "Phe177Ser mutation site (Claim 1)"
FT	Misc_difference	439
FT		/note= "Ile439Val mutation site (Claim 1)"

PN WO9801549-A2.
PD 15-JAN-1998.
PF 04-JUL-1997; CA0475.
PR 02-JAN-1997; US-034590.
PR 05-JUL-1996; US-021673.
PR 12-JUL-1996; US-021700.
PR 08-NOV-1996; US-029895.
PA (HSCR-) HSC RES & DEV LP.
PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
PI Fraser PE, Rommens JM, St George-Hyslop PH;
DR WPI; 98-286355/25.
DR N-PSDB; V04668.
PT New isolated mutant presenilin-1 genes - useful for developing
PT products for use in detection, diagnosis and therapy of Alzheimer's
PT disease and for drug screening
PS Disclosure; Page 199-200; 238pp; English.
CC This polypeptide comprises the murine presenilin-1 (PS1) homologue.
CC Its amino acid sequence was deduced from an isolated cDNA clone
CC (see V04668). Mutations in the human PS1 and PS2 genes (see
CC V04666-68) have been linked to the development in humans of forms
CC of familial Alzheimer's disease (FAD). All amino acids that are
CC mutated in analysed FAD pedigrees (see W23964) were conserved in
CC the murine homologue. Use of the nucleic acids and proteins
CC comprising or derived from presenilins can be made in screening and
CC diagnosing FAD, identifying and developing therapeutics for
CC treatment of FAD, and in producing cell lines and transgenic
CC animals useful as models of FAD. Methods for identifying
CC substances that bind to, or modulate the activity of a presenilin
CC protein, and methods for identifying substances that affect the
CC interaction of a presenilin-interacting protein with a presenilin
CC protein are also disclosed.
SQ Sequence 467 AA;

Query Match 74.4%; Score 1431.5; DB 1; Length 467;
Best Local Similarity 71.2%; Pred. No. 4e-147;
Matches 284; Conservative 36; Mismatches 50; Indels 29; Gaps 4;

```
Qy      1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
          |||||||||||||||||||||:||||||||| ||| |:|||||||||||| :|||| |
Db     71 EELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDTETVGQRAL 130

Qy     61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
          :|:|| |||||||:||| |||||||||| || |||:|||:||| |::|||||| ||||
Db    131 HSILNAAIMISVIVIMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYLGEVFKTYN 190

Qy    121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
          ||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||||||:||:||
Db    191 VAVDYVTVALLIWNFGVVGMIAIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLIL 250

Qy    181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240
            |||||||||||||||||||||||||| :|||||||| ||| | ||: || :|   :
Db    251 AVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGDPEAQ--RR 308

Qy    241 LPYDPEME--------EDSYDSFGEPSYPEVFEPPLTGYPG-----------EEL---- 277
          :| :|:          :||     :  : | :|     : |            :||
Db    309 VPKNPKYNTQRAERETQDSGSGNDDGGFSEEWEAQRDSHLGPHRSTPESRAAVQELSGSI 368

Qy    277 ---EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVF 333
             |: ||||||||||||||||||||:|| |||||||:||||||||||||||||||:|
Db    369 LTSEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIF 428

Qy    334 KKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
          ||||||||||||||:|||:|| ||:|||| || || ||
Db    429 KKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467
```

RESULT 12
W11839
ID W11839 standard; Protein; 467 AA.
AC W11839;
DT 07-MAY-1997 (first entry)
DE Human early onset Alzheimer's disease (EOAD) polypeptide.
KW Early onset Alzheimer's disease; EOAD; neurodegenerative disease;
KW diagnosis; therapy; inhibitor; antagonist; antibody.
OS Homo sapiens.

FH	Key	Location/Qualifiers
FT	misc_difference	26. .29
FT		/note= "unidentified amino acid residues"

```
CC   alternate splicing of the genomic DNA sequence. W05762 represents the
CC   coding sequence for wild type human PS-2. The presenilins are a family of
CC   highly conserved integral membrane proteins with a common structural
CC   motif, common alternate splicing patterns, and common mutational hot spot
CC   regions. Mutations in PS genes are implicated in familial Alzheimer's
CC   disease (AD) and possibly other diseases such as cerebral haemorrhage,
CC   schizophrenia, depression etc., so detection of mutations in the DNA
CC   encoding these sequences can be used for diagnosis of these diseases.
CC   These proteins, or vectors that express them or containing antisense
CC   sequences, antibodies selective for mutant forms of these proteins (such
CC   as W05736) and modulators of PS gene expression are potentially useful
CC   for treatment of AD etc. Transgenic animals are useful as models for drug
CC   screening. The antibodies can also be used e.g. for affinity purification
CC   and in immunoassays.
SQ   Sequence   467 AA;

  Query Match             74.3%;  Score 1429.5;  DB 1;  Length 467;
  Best Local Similarity   72.3%;  Pred. No. 6.6e-147;
  Matches  287;  Conservative   33;  Mismatches   52;  Indels   25;  Gaps    5;

Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
Db      71 EELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDTETVGQRAL 130

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
Db     131 HSILNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYLGEVFKTYN 190

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
Db     191 VAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLIL 250

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGAL- 240
Db     251 AVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVS 310

Qy     240 -QLPYDPE-MEEDSYDSFGE---PSYPEVFEPPLTGYPG------------EEL------ 277
Db     311 KNSKYNAESTERESQDTVAENDDGGFSEEWEAQRDSHLGPHRSTPESRAAVQELSSSILA 370

Qy     277 -EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKK 335
Db     371 GEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKK 430

Qy     336 ALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
Db     431 ALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467


RESULT  15
W28507
ID   W28507 standard; Protein; 407 AA.
AC   W28507;
DT   07-DEC-1997  (first entry)
DE   Partial AD3 sequence.
KW   AD3; AD4/AD3LP; Alzheimer's disease; chromosome; missegregation;
KW   presenilin; inhibitor; AD; trisomy 21.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   misc_difference 86
FT                   /label= mutation
FT                   /note= "M -> L"
FT   misc_difference 103
FT                   /label= mutation
FT                   /note= "H -> R"
FT   misc_difference 186
FT                   /label= mutation
FT                   /note= "A -> E"
FT   misc_difference 226
FT                   /label= mutation
FT                   /note= "L -> V"
FT   misc_difference 350
FT                   /label= mutation
FT                   /note= "C -> Y"
PN   WO9707213-A2.
PD   27-FEB-1997.
PF   15-AUG-1996; U13314.
PR   16-AUG-1995; US-002448.
PA   (HARD ) HARVARD COLLEGE.
PI   Li J,  Potter H;
DR   WPI; 97-165297/15.
DR   N-PSDB; T87402.
PT   Identifying genes which cause chromosome missegregation - useful for
PT   identifying causes of and treatments for diseases, e.g. Alzheimer's
PT   disease, cancer and ageing
PS   Disclosure; Fig 1; 77pp; English.
CC   Identifying genes which cause improper chromosome segregation,
CC   screening for inhibitors of chromosome missegregation and processes
CC   caused by genes encoding chromosome missegregation promoters
CC   was exemplified using Alzheimer's disease.  The sequences
CC   given in T87401 to T87426 can be used in the above methods.
CC   The five mutations indicated in the Features Table cosegregate
CC   with early-onset familial Alzheimer's disease. It is predicted
CC   that these mutations result in increased levels of cells with
CC   trisomy 21 in carriers of the mutation compared with non-carriers.
SQ   Sequence   407 AA;

  Query Match             74.3%;  Score 1429.5;  DB 1;  Length 407;
  Best Local Similarity   72.3%;  Pred. No. 5.4e-147;
  Matches  287;  Conservative   33;  Mismatches   52;  Indels   25;  Gaps    5;

Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
Db      11 EELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDTETVGQRAL 70

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
Db      71 HSILNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYLGEVFKTYN 130

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
Db     131 VAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLIL 190

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGAL- 240
Db     191 AVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVS 250

Qy     240 -QLPYDPE-MEEDSYDSFGE---PSYPEVFEPPLTGYPG------------EEL------ 277
Db     251 KNSKYNAESTERESQDTVAENDDGGFSEEWEAQRDSHLGPHRSTPESRAAVQELSSSILA 310

Qy     277 -EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKK 335
Db     311 GEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKK 370

Qy     336 ALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
Db     371 ALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 407


Search completed: March 20, 2000, 05:31:21
Job time: 4210 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 18, 2000, 19:55:31 ; Search time 26.47 Seconds
(without alignments)
186.879 Million cell updates/sec

Title: US-08-509-359B-138
Perfect score: 1923
Sequence: 1 EELTLKYGAKHVIMLFVPVT.....STDNLVRPFMDTLASHQLYI 372

Scoring table: BLOSUM62

Searched: 133990 seqs, 13297546 residues

Database : Issued_Patents_AA:*

Word size : 0

Number of hits that pass the threshold : 133990
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS9_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1923	100.0	448	2	US-08-967-101-137	Sequence 137, App
2	1923	100.0	372	2	US-08-967-101-138	Sequence 138, App
3	1923	100.0	448	2	US-08-592-541-137	Sequence 137, App
4	1923	100.0	372	2	US-08-592-541-138	Sequence 138, App
5	1907.5	99.2	447	2	US-08-875-972-29	Sequence 29, Appl
6	1591.5	82.8	376	2	US-08-875-972-2	Sequence 2, Appli
7	1429.5	74.3	467	2	US-08-967-101-134	Sequence 134, App
8	1429.5	74.3	407	2	US-08-875-972-4	Sequence 4, Appli
9	1429.5	74.3	467	2	US-08-592-541-134	Sequence 134, App
10	1429.5	74.3	467	3	US-08-670-964-2	Sequence 2, Appli
11	1429.5	74.3	463	3	US-08-670-964-4	Sequence 4, Appli
12	1423.5	74.0	467	2	US-08-967-101-2	Sequence 2, Appli
13	1423.5	74.0	467	2	US-08-592-541-2	Sequence 2, Appli
14	1412.5	73.5	463	2	US-08-670-479-18	Sequence 18, Appl
15	1381.5	71.8	467	2	US-08-967-101-4	Sequence 4, Appli
16	1381.5	71.8	467	2	US-08-592-541-4	Sequence 4, Appli
17	1127.5	58.6	541	2	US-08-967-101-166	Sequence 166, App
18	1127.5	58.6	541	2	US-08-592-541-166	Sequence 166, App
19	100	5.2	1294	2	US-08-819-288-3	Sequence 3, Appli
20	98	5.1	1321	1	US-08-261-822A-3	Sequence 3, Appli
21	98	5.1	1321	4	PCT-US95-07744A-3	Sequence 3, Appli
22	97	5.0	1334	2	US-08-996-545-2	Sequence 2, Appli
23	89	4.6	509	2	US-09-031-392-6	Sequence 6, Appli
24	88.5	4.6	452	1	US-08-117-361C-1	Sequence 1, Appli
25	88.5	4.6	3169	2	US-08-477-451-6	Sequence 6, Appli
26	87	4.5	492	2	US-08-355-844-3	Sequence 3, Appli
27	87	4.5	492	4	PCT-US95-16126-3	Sequence 3, Appli
28	82	4.3	502	1	US-08-484-840-3	Sequence 3, Appli
29	82	4.3	1873	1	US-08-336-257A-7	Sequence 7, Appli
30	82	4.3	502	1	US-08-483-094-3	Sequence 3, Appli
31	82	4.3	413	2	US-08-808-793-25	Sequence 25, Appl
32	81	4.2	2100	2	US-08-808-793-23	Sequence 23, Appl
33	80	4.2	837	1	US-07-923-976-2	Sequence 2, Appli
34	80	4.2	602	1	US-08-295-814E-2	Sequence 2, Appli
35	80	4.2	602	1	US-08-291-299-9	Sequence 9, Appli
36	80	4.2	438	2	US-08-846-762-8	Sequence 8, Appli
37	80	4.2	602	4	PCT-US93-01959-2	Sequence 2, Appli
38	80	4.2	602	4	PCT-US95-10579-9	Sequence 9, Appli
39	79.5	4.1	381	2	US-08-687-355A-6	Sequence 6, Appli
40	79	4.1	1873	1	US-08-435-675B-4	Sequence 4, Appli
41	79	4.1	1872	5	5386025-6	Patent No. 5386025
42	78.5	4.1	361	1	US-08-415-751-7	Sequence 7, Appli
43	78.5	4.1	381	2	US-08-687-355A-4	Sequence 4, Appli
44	78	4.1	466	1	US-08-194-338-11	Sequence 11, Appl
45	78	4.1	384	1	US-08-232-144-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT 1
US-08-967-101-137
; Sequence 137, Application US/08967101
; Patent No. 5840540
; GENERAL INFORMATION:
;   APPLICANT: ST. GEORGE-HYSLOP, PETER H
;   APPLICANT: ROMMENS, JOHANNA M
;   APPLICANT: FRASER, PAUL E
;   TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
;   TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
;   NUMBER OF SEQUENCES: 183
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: TESTA, HURWITZ & THIBEAULT
;     STREET: High Street Tower - 125 High Street
;     CITY: Boston
;     STATE: Massachusetts
;     COUNTRY: U.S.A.
;     ZIP: 02110
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/967,101
;     FILING DATE: 10-NOV-1997
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/592,541
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Pitcher, Edmund R.
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (617) 248-7000
;     TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 137:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 448 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-967-101-137


  Query Match            100.0%;  Score 1923;  DB 2;  Length 448;
  Best Local Similarity  100.0%;  Pred. No. 2.2e-190;
  Matches  372;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 136

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 196
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 136

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 196

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     197 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 256

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     257 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 316

Qy     241 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     317 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 376

Qy     301 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     377 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 436

Qy     361 FMDTLASHQLYI 372
                 ||||||||||||
Db     437 FMDTLASHQLYI 448


RESULT   4
US-08-592-541-138
; Sequence 138, Application US/08592541
; Patent No. 5986054
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  138:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  372 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-592-541-138


  Query Match           100.0%;  Score 1923;  DB 2;  Length 372;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-190;
  Matches  372;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240

Qy     241 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 300

Qy     301 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 360

Qy     361 FMDTLASHQLYI 372
                 ||||||||||||
Db     361 FMDTLASHQLYI 372


RESULT   5
US-08-875-972-29
; Sequence 29, Application US/08875972
; Patent No. 5985564
;  GENERAL INFORMATION:
;    APPLICANT:  Huntington Potter and Jinhue Li
;    TITLE OF INVENTION:  ASSAY FOR IDENTIFYING GENES CAUSING
;    TITLE OF INVENTION:  CHROMOSOME NON-DISJUNCTION
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02173-4799
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/875,972
;      FILING DATE:  08-AUG-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/002,448
;      FILING DATE:  16-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan Esq., Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  HU95-03PA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (781) 861-6240
;      TELEFAX:  (781) 861-9540
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-875-972-29
```

```
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-967-101-134


  Query Match             74.3%;  Score 1429.5;  DB 2;  Length 467;
  Best Local Similarity   72.3%;  Pred. No. 2e-139;
  Matches  287;  Conservative   33;  Mismatches   52;  Indels   25;  Gaps    5;

Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
Db      71 EELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDTETVGQRAL 130

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
Db     131 HSILNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYLGEVFKTYN 190

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
Db     191 VAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLIL 250

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGAL- 240
Db     251 AVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVS 310

Qy     240 -QLPYDPE-MEEDSYDSFGE---PSYPEVFEPPLTGYPG------------EEL------ 277
Db     311 KNSKYNAESTERESQDTVAENDDGGFSEEWEAQRDSHLGPHRSTPESRAAVQELSSSILA 370

Qy     277 -EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKK 335
Db     371 GEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKK 430

Qy     336 ALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
Db     431 ALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467


RESULT   8
US-08-875-972-4
; Sequence 4, Application US/08875972
; Patent No. 5985564
;  GENERAL INFORMATION:
;    APPLICANT:  Huntington Potter and Jinhue Li
;    TITLE OF INVENTION:  ASSAY FOR IDENTIFYING GENES CAUSING
;    TITLE OF INVENTION:  CHROMOSOME NON-DISJUNCTION
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02173-4799
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/875,972
;      FILING DATE:  08-AUG-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/002,448
;      FILING DATE:  16-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan Esq., Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  HU95-03PA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (781) 861-6240
;      TELEFAX:  (781) 861-9540
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  407 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-875-972-4


  Query Match             74.3%;  Score 1429.5;  DB 2;  Length 407;
  Best Local Similarity   72.3%;  Pred. No. 1.6e-139;
  Matches  287;  Conservative   33;  Mismatches   52;  Indels   25;  Gaps    5;

Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
Db      11 EELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDTETVGQRAL 70

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
Db      71 HSILNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYLGEVFKTYN 130

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
Db     131 VAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLIL 190

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGAL- 240
Db     191 AVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVS 250

Qy     240 -QLPYDPE-MEEDSYDSFGE---PSYPEVFEPPLTGYPG------------EEL------ 277
Db     251 KNSKYNAESTERESQDTVAENDDGGFSEEWEAQRDSHLGPHRSTPESRAAVQELSSSILA 310

Qy     277 -EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKK 335
Db     311 GEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKK 370

Qy     336 ALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
Db     371 ALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 407


RESULT   9
US-08-592-541-134
; Sequence 134, Application US/08592541
; Patent No. 5986054
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
```

```
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/670,964
;      FILING DATE:  26-JUN-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/001,142
;      FILING DATE:  13-JUL-1995
;      APPLICATION NUMBER:  60/001,501
;      FILING DATE:  18-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Han, William T
;      REGISTRATION NUMBER:  34,344
;      REFERENCE/DOCKET NUMBER:  P50358
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5219
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  463 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-670-964-4


  Query Match             74.3%;  Score 1429.5;  DB 3;  Length 463;
  Best Local Similarity   72.3%;  Pred. No. 1.9e-139;
  Matches  287;  Conservative   33;  Mismatches   52;  Indels   25;  Gaps    5;

Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           |||||||||||||||||||||||||:|||||||| ||| |:|||||||||||| :|||| |
Db      67 EELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDTETVGQRAL 126

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           :|:||  |||||||||| |||||||||| || |||:|||:||| |::|||||| ||||
Db     127 HSILNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYLGEVFKTYN 186

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
           ||:|| |: | :|||| |||: ||||||| |||||||||||||||||||||:||:||
Db     187 VAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLIL 246

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGAL- 240
             ||||||||||||||||||||||||||| :|||||||| ||| | ||: || :|  :
Db     247 AVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVS 306

Qy     240 -QLPYDPE-MEEDSYDSFGE---PSYPEVFEPPLTGYPG------------EEL------ 277
              |: |  | :| |:  |     : | :|     : |            :||
Db     307 KNSKYNAESTERESQDTVAENDDGGFSEEWEAQRDSHLGPHRSTPESRAAVQELSSSILA 366

Qy     277 -EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKK 335
            |: |||||||||||||||||||:|| |||||||:||||||||||||||||:|||
Db     367 GEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKK 426

Qy     336 ALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
           |||||||||||||:|||:|| ||:|||| || || ||
Db     427 ALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 463


RESULT  12
US-08-967-101-2
; Sequence 2, Application US/08967101
; Patent No. 5840540
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/967,101
;      FILING DATE:  10-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  467 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-967-101-2


  Query Match             74.0%;  Score 1423.5;  DB 2;  Length 467;
  Best Local Similarity   72.0%;  Pred. No. 8.2e-139;
  Matches  286;  Conservative   33;  Mismatches   53;  Indels   25;  Gaps    5;

Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           |||||||||||||||||||||||:|||||||| ||| |:|||||||||||| :|||| |
Db      71 EELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDTETVGQRAL 130

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           :|:||  |||||||||| |||||||||| || |||:|||:||| |::|||||| ||||
Db     131 HSILNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYLGEVFKTYN 190

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
           ||:|| |: | :|| | |||: ||||||| |||||||||||||||||||||:||:||
Db     191 VAVDYITVALLIWNLGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLIL 250

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGAL- 240
             ||||||||||||||||||||||||||| :|||||||| ||| | ||: || :|  :
Db     251 AVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVS 310

Qy     240 -QLPYDPE-MEEDSYDSFGE---PSYPEVFEPPLTGYPG------------EEL------ 277
              |: |  | :| |:  |     : | :|     : |            :||
Db     311 KNSKYNAESTERESQDTVAENDDGGFSEEWEAQRDSHLGPHRSTPESRAAVQELSSSILA 370

Qy     277 -EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKK 335
            |: |||||||||||||||||||:|| |||||||:||||||||||||||||:|||
Db     371 GEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKK 430

Qy     336 ALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
           |||||||||||||:|||:|| ||:|||| || || ||
Db     431 ALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467


RESULT  13
US-08-592-541-2
; Sequence 2, Application US/08592541
; Patent No. 5986054
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
```

```
          |: |||||||||||||||||||||||:|| |||||||:||||||||||||||||||:|||
Db    367 GEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKK 426

Qy    336 ALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
          |||||||||||||:|||:|| ||:|||| || || ||
Db    427 ALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 463


RESULT  15
US-08-967-101-4
; Sequence 4, Application US/08967101
; Patent No. 5840540
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/967,101
;      FILING DATE:  10-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  467 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-967-101-4


  Query Match           71.8%;  Score 1381.5;  DB 2;  Length 467;
  Best Local Similarity  68.7%;  Pred. No. 1.8e-134;
  Matches  274;  Conservative   38;  Mismatches   58;  Indels   29;  Gaps    4;

Qy      1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
          |||||||||||||||||||||||:||||||||| ||| |:|||||||||||| :|||| |
Db     71 EELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDTETVGQRAL 130

Qy     61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
          :|:||  |||||||:||| |||||||||| || |||:|||:||| |::|||||| ||||
Db    131 HSILNAAIMISVIVIMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYLGEVFKTYN 190

Qy    121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
          | :|| |: | :||:| |||: ||||||| ||||||||||||||||||||||:||:||
Db    191 VXVDYVTVALLIWNWGVVGMIAIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLIL 250

Qy    181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240
            |||||||||||||||||||||||||| :|||||||| ||| | ||: || :|   :
Db    251 AVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGDPEAQ--RR 308

Qy    241 LPYDPEME--------EDSYDSFGEPSYPEVFEPPLTGYPG------------EEL---- 277
          :| :|:          :||     :  : | :|     : |            :||
Db    309 VPKNPKYNTQRAERETQDSGSGNDDGGFSEEWEAQRDSHLGPHRSTPESRAAVQELSGSI 368

Qy    277 ---EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVF 333
              |: |||||||||||||||||||||:|| |||||||:|| ||||||||| |||||::
Db    369 LTSEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACXVAILIGLCLXLLLLAIY 428

Qy    334 KKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
          ||  || ||||||| :| |:|| ||:|||| || || ||
Db    429 KKGXPAXPISITFGFVFXFATDYLVQPFMDQLAFHQFYI 467


Search completed: March 18, 2000, 19:55:32
Job time: 3230 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 18, 2000, 14:11:52 ; Search time 41.25 Seconds
(without alignments)
425.381 Million cell updates/sec

Title: US-08-509-359B-138
Perfect score: 1923
Sequence: 1 EELTLKYGAKHVIMLFVPVT.....STDNLVRPFMDTLASHQLYI 372

Scoring table: BLOSUM62

Searched: 142080 seqs, 47169319 residues

Database : PIR_62:*

Word size : 0

Number of hits that pass the threshold : 142080
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1923	100.0	448	2	I58098	E5-1 protein - hum
2	1915	99.6	448	2	A56993	presenilin 2 - hum
3	1802.5	93.7	442	2	I39174	seven trans-membra
4	1534	79.8	449	2	JC5391	presenilin-beta -
5	1431.5	74.4	467	2	I78388	S182 protein - mou
6	1431.5	74.4	433	2	JC5390	presenilin-alpha -
7	1429.5	74.3	467	2	S58396	presenilin 1, spli
8	1429.5	74.3	463	2	S63683	presenilin I-463 -
9	1413.5	73.5	463	2	JC5081	presenilin 1 prote
10	1413.5	73.5	467	2	JC5080	presenilin 1 prote
11	1011	52.6	374	2	S63684	presenilin 1, spli
12	956	49.7	461	2	S60253	sel-12 protein - C
13	524.5	27.3	358	2	T15184	hypothetical prote
14	513	26.7	453	2	T00724	presenilin homolog
15	274	14.2	465	2	A43459	sperm membrane pro
16	110	5.7	2016	2	A38195	sodium channel pro
17	104	5.4	826	2	T02268	potassium transpor
18	104	5.4	398	2	H75043	mg2+ transport pro
19	103	5.4	382	2	S47882	ubiquinol--cytochr
20	102	5.3	379	2	I48135	ubiquinol--cytochr
21	101.5	5.3	1681	2	A55138	sodium channel mNa
22	100.5	5.2	324	2	S36646	integrin-associate
23	99.5	5.2	1840	1	CHRTM1	sodium channel pro
24	99.5	5.2	379	2	JC6178	serotonin receptor
25	99	5.1	2019	2	A33996	sodium channel pro
26	98.5	5.1	531	2	T11596	hypothetical prote
27	98	5.1	461	2	T11829	NADH dehydrogenase
28	98	5.1	217	2	S01095	hypothetical prote
29	97	5.0	381	2	T11440	ubiquinol--cytochr
30	97	5.0	447	2	S52968	NADH dehydrogenase
31	97	5.0	238	2	S02063	H+-transporting AT
32	97	5.0	441	2	S13425	endothelin recepto
33	97	5.0	299	2	D65187	hypothetical 33.7
34	96.5	5.0	308	2	S22928	ubiquinol--cytochr
35	96.5	5.0	1070	2	S25834	rad3 protein - fis
36	96	5.0	507	2	B64433	probable O-antigen
37	96	5.0	592	2	E70488	cytochrome-c oxida
38	95.5	5.0	488	1	QXASM4	NADH dehydrogenase
39	95.5	5.0	442	1	JQ1042	endothelin recepto
40	95.5	5.0	294	2	D53290	oligopeptide trans
41	95.5	5.0	420	2	A47649	probable inner mem
42	95.5	5.0	768	2	S52684	probable membrane
43	95.5	5.0	799	2	H71255	probable cell divi
44	95	4.9	492	2	A30797	glucose transport
45	95	4.9	484	2	S75022	bacteriochlorophyl

ALIGNMENTS

RESULT 1
I58098
E5-1 protein - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C;Accession: I58098
R;Rogaev, E.I.; Sherrington, R.; Rogaeva, E.A.; Levesque, G.; Ikeda, M.; Liang, Y.; C
.; Cohen, D.; Lannfelt, L.; Fraser, P.E.; Rommens, J.M.; St George-Hyslop, P.H.
Nature 376, 775-778, 1995
A;Title: Familial Alzheimer's disease in kindreds with missense mutations in a gene o
A;Reference number: I58098; MUID:95379971
A;Accession: I58098
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-448 <RES>
A;Cross-references: GB:L44577; NID:g950347; PIDN:AAC42012.1; PID:g950348
C;Genetics:
A;Gene: E5-1
C;Superfamily: presenilin

Query Match 100.0%; Score 1923; DB 2; Length 448;
Best Local Similarity 100.0%; Pred. No. 3.5e-138;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 136

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 196

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     197 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 256

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     257 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 316

Qy     241 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     317 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 376

Qy     301 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     377 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 436

Qy     361 FMDTLASHQLYI 372
           ||||||||||||
Db     437 FMDTLASHQLYI 448
```

RESULT 2
A56993
presenilin 2 - human
N;Alternate names: Alzheimer's disease protein 4

```
Qy    240 -QLPY----DPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFY 294
          |: :     ||    : :   |           |       : :|||||||||||||||
Db    319 QQVQHIDRNTPEGANSTVEDAAETRIQ-------TQSNLSSEDPDEERGVKLGLGDFIFY 371

Qy    295 SVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFST 354
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    372 SVLVGKAAATASGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFST 431

Qy    355 DNLVRPFMDTLASHQLYI 372
          ||:|||||||||||||:||
Db    432 DNIVRPFMDTLASHQMYI 449


RESULT   5
I78388
S182 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 29-Sep-1999
C;Accession: I78388
R;Sherrington, R.; Rogaev, E.I.; Liang, Y.; Rogaeva, E.A.; Levesque, G.; Ikeda, M.; Chi,
ero, I.; Pinessi, L.; Nee, L.; Chumakov, I.; Pollen, D.; Brookes, A.; Sanseau, P.; Polin
E.; Rommens, J.M.; St George-Hyslop, P.H.
Nature 375, 754-760, 1995
A;Title: Cloning of a gene bearing missense mutations in early-onset familial Alzheimer'
A;Reference number: I58095; MUID:95319502
A;Accession: I78388
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-467 <RES>
A;Cross-references: GB:L42177; NID:g904129; PIDN:AAC42094.1; PID:g904130
C;Superfamily: presenilin


  Query Match             74.4%;  Score 1431.5;  DB 2;  Length 467;
  Best Local Similarity   71.2%;  Pred. No. 5.1e-101;
  Matches  284;  Conservative   36;  Mismatches   50;  Indels   29;  Gaps    4;

Qy      1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
          ||||||||||||||||||||||:|||||||| ||| |:|||||||||| :|||| |
Db     71 EELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDTETVGQRAL 130

Qy     61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
          :|:||  |||||||:||| |||||||||| || |||:|||:||| |::|||||| ||||
Db    131 HSILNAAIMISVIVIMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYLGEVFKTYN 190

Qy    121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
          ||:|| |: | :|||| |||: ||||||| |||||||||||||||||||||:||:||
Db    191 VAVDYVTVALLIWNFGVVGMIAIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLIL 250

Qy    181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240
            |||||||||||||||||||||||||||| :|||||||| ||| | ||: || :|   :
Db    251 AVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGDPEAQ--RR 308

Qy    241 LPYDPEME--------EDSYDSFGEPSYPEVFEPPLTGYPG------------EEL---- 277
          :| :|:          :||     :  : | :|     : |            :||
Db    309 VPKNPKYNTQRAERETQDSGSGNDDGGFSEEWEAQRDSHLGPHRSTPESRAAVQELSGSI 368

Qy    277 ---EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVF 333
             |: |||||||||||||||||||||:|| |||||||:|||||||||||||||||:|
Db    369 LTSEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIF 428

Qy    334 KKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
          |||||||||||||:|||:|| ||:|||| || || ||
Db    429 KKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467


RESULT   6
JC5390
presenilin-alpha - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 04-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 29-Sep-1999
C;Accession: JC5390
R;Tsujimura, A.; Yasojima, K.; Hashimoto-Gotoh, T.
Biochem. Biophys. Res. Commun. 231, 392-396, 1997
A;Title: Cloning of Xenopus presenilin-alpha and -beta cDNAs and their differential e
A;Reference number: JC5390; MUID:97223465
A;Accession: JC5390
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-433 <TSU>
A;Cross-references: DDBJ:D84427; NID:g1944353; PIDN:BAA19570.1; PID:d1020347; PID:g19
A;Experimental source: brain
C;Comment: This protein plays a role in negative regulation of apoptotic cascades dur
C;Superfamily: presenilin
F;48-66/Domain: transmembrane #status predicted <TM1>
F;99-119/Domain: transmembrane #status predicted <TM2>
F;130-149/Domain: transmembrane #status predicted <TM3>
F;161-178/Domain: transmembrane #status predicted <TM4>
F;187-203/Domain: transmembrane #status predicted <TM5>
F;210-227/Domain: transmembrane #status predicted <TM6>
F;374-394/Domain: transmembrane #status predicted <TM7>


  Query Match             74.4%;  Score 1431.5;  DB 2;  Length 433;
  Best Local Similarity   72.0%;  Pred. No. 4.7e-101;
  Matches  286;  Conservative   27;  Mismatches   59;  Indels   25;  Gaps    4;

Qy      1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
          ||||||||||||||||||||||||:||||||||| |||  :|||||||||||| ||||| |
Db     37 EELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRFDGQLIYTPFTEDTESVGQRAL 96

Qy     61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
          ||:||  ||||||:|||| |||||||||| ||||||:|||:||| |:|||||||| ||||
Db     97 NSILNATIMISVIIVMTILLVVLYKYRCYKVIHGWLIISSLLLLFFFSYIYLGEVFKTYN 156

Qy    121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
          ||:|| || | :|||| |||:||||||||:|||||||||||||||||||||: |:||
Db    157 VAVDYITLALLIWNFGVVGMICIHWKGPLLLQQAYLIMISALMALVFIKYLPEWTTWLIL 216

Qy    181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPS-SQGAL 239
            |||||||||| |||||||||||||||| :|||||||| |:| | ||  ||   | |
Db    217 AVISVYDLVAVLSPKGPLRMLVETAQERNETLFPALIYSSTMIWLVNMADGDPGLKQSAS 276

Qy    240 QLPYDPEM-----EEDSYDSFGEPSYPEVFEP-------PLTGYPGEEL----------- 277
            |: :        ||  |     :  :|        |:   |   :
Db    277 TKTYNTQAPTAHPRSDSAASDDNGGFDTTWEDHRNAQIGPINSTPESRVAVQALPSNSPP 336

Qy    277 -EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKK 335
           |: |||||||||||||||||||:|| |||||||||||||||||||||||||:||
Db    337 SEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTLACFVAILIGLCLTLLLLAIFKK 396

Qy    336 ALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
          |||||||||||||:|||:|| ||:|||| || || ||
Db    397 ALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 433


RESULT   7
S58396
presenilin 1, splice form 467 - human
N;Alternate names: Alzheimer's disease protein 3; protein S182
C;Species: Homo sapiens (man)
C;Date: 29-Jan-1998 #sequence_revision 13-Feb-1998 #text_change 29-Sep-1999
C;Accession: S58396; S71401; S71402
R;Sherrington, R.; Rogaev, E.I.; Liang, Y.; Rogaeva, E.A.; Levesque, G.; Ikeda, M.; C
ero, I.; Pinessi, L.; Nee, L.; Chumakov, I.; Pollen, D.; Brookes, A.; Sanseau, P.; Po
E.; Rommens, J.M.; St George-Hyslop, P.H.
Nature 375, 754-760, 1995
A;Title: Cloning of a gene bearing missense mutations in early-onset familial Alzheim
A;Reference number: I58095; MUID:95319502
A;Accession: S58396
A;Molecule type: mRNA
A;Residues: 1-467 <SHE>
A;Cross-references: EMBL:L42110; NID:g904118; PIDN:AAB46416.1; PID:g904119
R;Vidal, R.; Ghiso, J.; Wisniewski, T.; Frangione, B.
FEBS Lett. 393, 19-23, 1996
A;Title: Alzheimer's presenilin 1 gene expression in platelets and megakaryocytes. Id
```

```
Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           :||||  |||||||||||| |||||||||||| || |||:|||:||| |::|||||| ||||
Db     127 HSVLNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYLGEVFKTYN 186

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
           ||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||||:||:||
Db     187 VAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLIL 246

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGAL- 240
           ||||||||||||||||||||||||||| :|||||||| ||| | ||: || :|  :
Db     247 AVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVS 306

Qy     240 -QLPYD------------PEMEEDSYDSFGEPSYPEVFEP----PLTGYPGEEL------ 277
            |:              || ::  :    |        |      ::    :|:
Db     307 KNTKYNAQGTEREAQASVPENDDGGFSEEWEAQRDSQLGPHRSTSVSRAAVQEISSSIPA 366

Qy     277 -EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKK 335
            |: |||||||||||:||||||||:|| |||||||:|||||||||||||||||:|||
Db     367 SEDPEERGVKLGLGDFVFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKK 426

Qy     336 ALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
           |||||||||||||:|||:|| ||:|||| || || ||
Db     427 ALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 463
```

RESULT 10
JC5080
presenilin 1 protein isoform 467 - lesser mouse lemur
C;Species: Microcebus murinus (lesser mouse lemur)
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 29-Sep-1999
C;Accession: JC5080
R;Calenda, A.; Mestre-Frances, N.; Czech, C.; Pradier, L.; Petter, A.; Bons, N.; Bellis,
Biochem. Biophys. Res. Commun. 228, 430-439, 1996
A;Title: Molecular cloning, sequencing, and brain expression of the presenilin 1 gene in
A;Reference number: JC5080; MUID:97079199
A;Accession: JC5080
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-467 <CAL>
A;Cross-references: EMBL:Z71333; NID:g1707591; PIDN:CAA95930.1; PID:e248527; PID:g170759
A;Experimental source: brain
C;Comment: This protein is an intergral membrane protein with seven transmembrane domain
C;Genetics:
A;Gene: ps1
A;Map position: 14
C;Superfamily: presenilin
C;Keywords: transmembrane protein
F;82-100/Domain: transmembrane #status predicted <TM1>
F;133-154/Domain: transmembrane #status predicted <TM2>
F;164-185/Domain: transmembrane #status predicted <TM3>
F;195-213/Domain: transmembrane #status predicted <TM4>
F;221-238/Domain: transmembrane #status predicted <TM5>
F;244-261/Domain: transmembrane #status predicted <TM6>
F;408-428/Domain: transmembrane #status predicted <TM7>

  Query Match   73.5%; Score 1413.5; DB 2; Length 467;
  Best Local Similarity 71.0%; Pred. No. 1.2e-99;
  Matches 282; Conservative   34; Mismatches   56; Indels   25; Gaps   4;

```
Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           |||||||||||||||||||||:||||||||| ||| |:|||||||||||| :|||| |
Db      71 EELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDTETVGQRAL 130

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           :||||  |||||||||||| |||||||||||| || |||:|||:||| |::|||||| ||||
Db     131 HSVLNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYLGEVFKTYN 190

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
           ||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||||:||:||
Db     191 VAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLIL 250
```

```
Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGAL- 240
           ||||||||||||||||||||||||||| :|||||||| ||| | ||: || :|  :
Db     251 AVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVS 310

Qy     240 -QLPYD------------PEMEEDSYDSFGEPSYPEVFEP----PLTGYPGEEL------ 277
            |:              || ::  :    |        |      ::    :|:
Db     311 KNTKYNAQGTEREAQASVPENDDGGFSEEWEAQRDSQLGPHRSTSVSRAAVQEISSSIPA 370

Qy     277 -EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKK 335
            |: |||||||||||:||||||||:|| |||||||:|||||||||||||||||:|||
Db     371 SEDPEERGVKLGLGDFVFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKK 430

Qy     336 ALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
           |||||||||||||:|||:|| ||:|||| || || ||
Db     431 ALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467
```

RESULT 11
S63684
presenilin 1, splice form 374 - human
N;Alternate names: Alzheimer's disease protein 3
C;Species: Homo sapiens (man)
C;Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 29-Sep-1999
C;Accession: S63684
R;Sahara, N.; Yahagi, Y.; Takagi, H.; Kondo, T.; Okochi, M.; Usami, M.; Shirasawa, T.
FEBS Lett. 381, 7-11, 1996
A;Title: Identification and characterization of presenilin I-467, I-463 and I-374.
A;Reference number: S63683; MUID:96193901
A;Accession: S63684
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-374 <SAH>
A;Cross-references: EMBL:U40380; NID:g1244639; PIDN:AAB05895.1; PID:g1244640
C;Genetics:
A;Gene: GDB:PSEN1; AD3; FAD; S182; PS1
A;Cross-references: GDB:135682; OMIM:104311
A;Map position: 14q24.3-14q24.3
C;Superfamily: presenilin

  Query Match   52.6%; Score 1011; DB 2; Length 374;
  Best Local Similarity 81.8%; Pred. No. 2.4e-69;
  Matches 193; Conservative   18; Mismatches   25; Indels   0; Gaps   0;

```
Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           |||||||||||||||||||||:||||||||| ||| |:|||||||||||| :|||| |
Db      67 EELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDTETVGQRAL 126

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           :|:||  |||||||||||| |||||||||||| || |||:|||:||| |::|||||| ||||
Db     127 HSILNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYLGEVFKTYN 186

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
           ||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||||:||:||
Db     187 VAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLIL 246

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQ 236
           |||||||||||||||||||||||||||| :|||||||| ||| | ||: || :|
Db     247 AVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGDPEAQ 302
```

RESULT 12
S60253
sel-12 protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 13-Sep-1998
C;Accession: S60253
R;Levitan, D.; Greenwald, I.
Nature 377, 351-354, 1995
A;Title: Facilitation of lin-12-mediated signalling by sel-12, a Caenorhabditis elega
A;Reference number: S60253; MUID:96032531
A;Accession: S60253
A;Status: preliminary; nucleic acid sequence not shown

```
          ||||||||||||||||:||     |  |  ||::||: ||  ||:||:|: :||||||||||
Db    369 KLGLGDFIFYSVLVGRAAMY---DLMTVYACYLAIISGLGCTLILLSVYNRALPALPISI 425

Qy    345 TFGLIFYFSTDNLVRPFM 362
            |::||| |  |: ||:
Db    426 MLGVVFYFLTRLLMEPFV 443


RESULT  15
A43459
sperm membrane protein spe-4 - Caenorhabditis elegans
N;Alternate names: probable integral membrane protein
C;Species: Caenorhabditis elegans
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 09-Sep-1997
C;Accession: A43459; S24632; S24633
R;L'Hernault, S.W.; Arduengo, P.M.
J. Cell Biol. 119, 55-68, 1992
A;Title: Mutation of a putative sperm membrane protein in Caenorhabditis elegans prevent
A;Reference number: A43459; MUID:92407040
A;Accession: A43459
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA; mRNA
A;Residues: 1-465 <LHE>
A;Cross-references: EMBL:Z14067; NID:g6868; PID:g6869; EMBL:Z14066; NID:g6870; PID:g6871
A;Experimental source: strain Bristol N2
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1992
C;Genetics:
A;Introns: 69/3; 154/3; 200/1; 224/3; 300/1; 386/1; 435/1


  Query Match             14.2%;  Score 274;  DB 2;  Length 465;
  Best Local Similarity   21.5%;  Pred. No. 1.5e-13;
  Matches   93;  Conservative   80;  Mismatches  127;  Indels  132;  Gaps   14;

Qy     38 EKNGQLIYTPFTEDT--PSVGQRLLNSVLN----TLIMISVIVVMTIFLVVLYKYRCYKF 91
          | | :|  | | : :   : |  ||:  :|     |::  |  :|  |  ||: :|  :
Db     42 EVNSELSKTYFLDPSFEQTTGNLLLDGFINGVGTILVLGCVSFIMLAF--VLFDFR--RI 97

Qy     92 IHGWLIMSSLMLLF---------LFTYIYLGEVLKTYNVAMDYPTLLL-----TVWNFGA 137
          :  || :| |::||         :|: ::  :    |     | |::|      |: ||
Db     98 VKAWLTLSCLLILFGVSAQTLHDMFSQVFDQDDNNQY-----YMTIVLIVVPTVVYGFG- 152

Qy    138 VGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVILGAISVYDLVAVLCPKGP 197
           |:        |:| | :::   :|::: :::  |  : | :|  :  :|| ||| | ||
Db    152 -GIYAFFSNSSLILHQIFVVTNCSLISVFYLRVFPSKTTWFVLWIVLFWDLFAVLAPMGP 210

Qy    198 LRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQG----------ALQLPYDPE 246
          |: : | | : :: :   :::|:        :  : :::|            ::     |
Db    211 LKKVQEKASDYSKCVLNLIMFSANEKRLTAGSNQEETNEGEESTIRRTVKQTIEYYTKRE 270

Qy    247 MEEDSY-------------DSF---------GEPSYPEVFEPPLTGYPGEELEEEEE--- 282
           ::| :              ||           |||  |: |   |    ||| ::|
Db    271 AQDDEFYQKIRQRRAAINPDSVPTEHSPLVEAEPSPIELKEKNST----EELSDDESDTS 326

Qy    282 ----------------------------------------------------------R 282

Db    327 ETSSGSSNLSSSDSSTTVSTSDISTAEECDQKEWDDLVSNSLPNNDKRPATAADALNDGE 386

Qy    283 GVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPI 342
          ::|| |||:|||:|:|:|||:|       ::  : || || :|| : :  :   ||||:
Db    387 VLRLGFGDFVFYSLLIGQAAASGCP--FAVISAALGILFGLVVTLTVFSTEESTTPALPL 444

Qy    343 SITFGLIFYFST 354
          :  |   |||:
Db    445 PVICGTFCYFSS 456


Search completed: March 18, 2000, 14:11:54
Job time: 491 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 18, 2000, 18:09:18 ; Search time 32.57 Seconds
(without alignments)
341.103 Million cell updates/sec

Title: US-08-509-359B-138
Perfect score: 1923
Sequence: 1 EELTLKYGAKHVIMLFVPVT.....STDNLVRPFMDTLASHQLYI 372

Scoring table: BLOSUM62

Searched: 82229 seqs, 29864866 residues

Database : SwissProt_38:*

Word size : 0

Number of hits that pass the threshold : 82229

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1923	100.0	448	1	PSN2_HUMAN	P49810 homo sapien
2	1876	97.6	448	1	PSN2_MOUSE	Q61144 mus musculu
3	1866	97.0	448	1	PSN2_RAT	O88777 rattus norv
4	1855	96.5	445	1	PSN2_MICMU	P79801 microcebus
5	1534	79.8	449	1	PSN2_XENLA	O12977 xenopus lae
6	1431.5	74.4	467	1	PSN1_MOUSE	P49769 mus musculu
7	1431.5	74.4	433	1	PSN1_XENLA	O12976 xenopus lae
8	1431	74.4	468	1	PSN1_RAT	P97887 rattus norv
9	1429.5	74.3	467	1	PSN1_HUMAN	P49768 homo sapien
10	1413.5	73.5	467	1	PSN1_MICMU	P79802 microcebus
11	1127.5	58.6	541	1	PSN_DROME	O02194 drosophila
12	1032	53.7	836	1	YL4K_CAEEL	Q20076 caenorhabdi
13	972	50.5	461	1	SE12_CAEEL	P52166 caenorhabdi
14	524.5	27.3	358	1	HOP1_CAEEL	O02100 caenorhabdi
15	513	26.7	453	1	PSNH_ARATH	O64668 arabidopsis
16	274	14.2	465	1	SPE4_CAEEL	Q01608 caenorhabdi
17	110	5.7	2016	1	CIN5_HUMAN	Q14524 homo sapien
18	108	5.6	381	1	CYB_DASCR	Q34302 dasycercus
19	103	5.4	382	1	CYB_DIDMA	P41303 didelphis m
20	103	5.4	380	1	CYB_MICLO	P56731 microtus lo
21	102.5	5.3	381	1	CYB_ANTFL	Q33706 antechinus
22	102	5.3	381	1	CYB_DASMA	O03522 dasyurus ma
23	102	5.3	381	1	CYB_PSENI	Q35553 pseudantech
24	101.5	5.3	381	1	CYB_NINYV	Q35196 ningaui yvo
25	101.5	5.3	381	1	CYB_PARAP	Q35377 parantechin
26	101	5.3	381	1	CYB_PLAMS	Q35533 planigale m
27	100.5	5.2	381	1	CYB_PSEMD	O03543 pseudantech
28	100	5.2	381	1	CYB_DASGE	O20604 dasyurus ge
29	99.5	5.2	1840	1	CIN4_RAT	P15390 rattus norv
30	99	5.1	2019	1	CIN5_RAT	P15389 rattus norv
31	99	5.1	381	1	CYB_ANTMI	O63534 antechinus
32	99	5.1	381	1	CYB_PHATA	Q35673 phascogale
33	98.5	5.1	531	1	YDFG_SCHPO	Q10487 schizosacch
34	98	5.1	381	1	CYB_ANTME	Q33782 antechinus
35	98	5.1	381	1	CYB_DASAL	Q34289 dasyurus al
36	98	5.1	381	1	CYB_SMICR	Q35810 sminthopsis
37	98	5.1	460	1	NU4M_GADMO	P55781 gadus morhu
38	98	5.1	217	1	YPRA_ECOLI	P13974 escherichia
39	97.5	5.1	381	1	CYB_DASHA	Q34321 dasyurus ha
40	97.5	5.1	381	1	CYB_PARBI	Q35380 parantechin
41	97	5.0	238	1	ATP6_BACP3	P09218 bacillus ps
42	97	5.0	381	1	CYB_DASVI	Q34399 dasyurus vi
43	97	5.0	381	1	CYB_MACRO	P92671 macropus ro
44	97	5.0	347	1	ETBR_COTJA	Q90328 coturnix co
45	97	5.0	447	1	NU4M_APILI	P34853 apis mellif

ALIGNMENTS

RESULT 1
PSN2_HUMAN
ID PSN2_HUMAN STANDARD; PRT; 448 AA.
AC P49810;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PRESENILIN 2 (PS-2) (STM-2) (E5-1) (AD3LP) (AD5).
GN PSEN2 OR PSNL2 OR AD4 OR PS2 OR STM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT FAD ILE-141.
RX MEDLINE; 95365816.
RA LEVY-LAHAD E., WASCO W., POORKAJ P., ROMANO D.M., OSHIMA J.,
RA PETTINGELL W.H., YU C.-E., JONDRO P.D., SCHMIDT S.D., WANG K.,
RA CROWLEY A.C., FU Y.-H., GUENETTE S.Y., GALAS D., NEMENS E.,
RA WIJSMAN E.M., BIRD T.D., SCHELLENBERG G.D., TANZI R.E.;
RT "Candidate gene for the chromosome 1 familial Alzheimer's disease
RT locus.";
RL Science 269:973-977(1995).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS FAD ILE-141 AND VAL-239.
RC TISSUE=BRAIN, AND COLON;
RX MEDLINE; 95379971.
RA ROGAEV E.I., SHERRINGTON R., ROGAEVA E.A., LEVESQUE G., IKEDA M.,
RA LIANG Y., CHI H., LIN C., HOLMAN K., TSUDA T., MAR L., SORBI S.,
RA NACMIAS B., PIACENTINI S., AMADUCCI L., CHUMAKOV I., COHEN D.,
RA LANNFELT L., FRASER P.E., ROMMENS J.M., ST GEORGE-HYSLOP P.H.;
RT "Familial Alzheimer's disease in kindreds with missense mutations in
RT a gene on chromosome 1 related to the Alzheimer's disease type 3
RT gene.";
RL Nature 376:775-778(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96109229.
RA LI J., MA J., POTTER H.;
RT "Identification and expression analysis of a potential familial
RT Alzheimer disease gene on chromosome 1 related to AD3.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:12180-12184(1995).
RN [4]
RP SEQUENCE FROM N.A.
RA LEVY-LAHAD E., POORKAJ P., WANG K., FU Y.H., OSHIMA J.,
RA MULLIGAN J., SCHELLENBERG G.D.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP REVIEW ON VARIANTS.
RX MEDLINE; 98180715.
RA CRUTS M., VAN BROECKHOVEN C.;
RT "Presenilin mutations in Alzheimer's disease.";
RL Hum. Mutat. 11:183-190(1998).
RN [6]
RP VARIANT AD HIS-62.
RX MEDLINE; 98046005.
RA CRUTS M., VAN DUIJN C.M., BACKHOVENS H., VAN DEN BROECK M.,
RA WEHNERT A., SERNEELS S., SHERRINGTON R., HUTTON M., HARDY J.,
RA ST GEORGE-HYSLOP P.H., HOFMAN A., VAN BROECKHOVEN C.;
RT "Estimation of the genetic contribution of presenilin-1 and -2
RT mutations in a population-based study of presenile Alzheimer
RT disease.";
RL Hum. Mol. Genet. 7:43-51(1998).

```
DR   PFAM; PF01080; Presenilin; 1.
KW   Transmembrane; Alternative initiation.
FT   CHAIN         1    448       PRESENILIN 2.
FT   CHAIN       298    448       PRESENILIN 2-SHORT.
FT   TRANSMEM     88    106       POTENTIAL.
FT   TRANSMEM    142    160       POTENTIAL.
FT   TRANSMEM    167    188       POTENTIAL.
FT   TRANSMEM    203    219       POTENTIAL.
FT   TRANSMEM    230    246       POTENTIAL.
FT   TRANSMEM    253    269       POTENTIAL.
FT   TRANSMEM    288    305       POTENTIAL.
FT   TRANSMEM    387    406       POTENTIAL.
FT   TRANSMEM    413    429       POTENTIAL.
FT   CONFLICT     87     87       R -> H (IN REF. 2).
FT   CONFLICT    226    226       A -> V (IN REF. 2).
FT   CONFLICT    324    324       MISSING (IN REF. 2).
SQ   SEQUENCE   448 AA;  49955 MW;  680ACF19 CRC32;


  Query Match             97.6%;  Score 1876;  DB 1;  Length 448;
  Best Local Similarity   97.3%;  Pred. No. 1.7e-121;
  Matches  362;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 EELTLKYGAKRVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 136

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 196

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
           |||||||| | ||||||||||||||||| |||||||:||||||||||||||||||||||||
Db     197 VAMDYPTLFLAVWNFGAVGMVCIHWKGPLALQQAYLIVISALMALVFIKYLPEWSAWVIL 256

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     257 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 316

Qy     241 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 300
           |||||||||||||||||||| || || ||||||||||||||||||||||||||||||||||
Db     317 LPYDPEMEEDSYDSFGEPSYPEAFEAPLPGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 376

Qy     301 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 360
           |||||:||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db     377 AAATGNGDWNTTLACFIAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 436

Qy     361 FMDTLASHQLYI 372
           ||||||||||||
Db     437 FMDTLASHQLYI 448


RESULT   3
PSN2_RAT
ID   PSN2_RAT       STANDARD;      PRT;   448 AA.
AC   O88777; O35546; O08947;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   PRESENILIN 2 (PS-2).
GN   PSEN2 OR PSNL2 OR PS2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=BRAIN;
RA   FRENTZEL S., ABDEL A.S., LUEBBERT H.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=BRAIN;
RX   MEDLINE; 97473536.
RA   TAKAHASHI H., MERCKEN M., NAKAZATO Y., NOGUCHI K., MURAYAMA M.,
RA   IMAHORI K., TAKASHIMA A.;
RT   "Cloning of cDNA and expression of the gene encoding rat
RT   presenilin-2.";
RL   Gene 197:383-387(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=BRAIN;
RX   MEDLINE; 98207716.
RA   TANAHASHI H., TABIRA T.;
RT   "Cloning of the cDNA encoding rat presenilin-2.";
RL   Biochim. Biophys. Acta 1396:259-262(1998).
CC   -!- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE
CC       EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE. MAY
CC       FUNCTION IN THE CYTOPLASMIC PARTITIONING OF PROTEINS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X99267; CAA67663.1; -.
DR   EMBL; D83700; BAA22832.1; -.
DR   EMBL; AB004454; BAA20406.1; -.
DR   PFAM; PF01080; Presenilin; 1.
KW   Transmembrane.
FT   TRANSMEM     88    106       POTENTIAL.
FT   TRANSMEM    142    160       POTENTIAL.
FT   TRANSMEM    167    188       POTENTIAL.
FT   TRANSMEM    203    219       POTENTIAL.
FT   TRANSMEM    230    246       POTENTIAL.
FT   TRANSMEM    253    269       POTENTIAL.
FT   TRANSMEM    288    305       POTENTIAL.
FT   TRANSMEM    387    406       POTENTIAL.
FT   TRANSMEM    413    429       POTENTIAL.
FT   CONFLICT      7      7       S -> T (IN REF. 1).
FT   CONFLICT     86     87       KH -> ND (IN REF. 3).
SQ   SEQUENCE   448 AA;  50051 MW;  EE0DF681 CRC32;


  Query Match             97.0%;  Score 1866;  DB 1;  Length 448;
  Best Local Similarity   96.8%;  Pred. No. 8.4e-121;
  Matches  360;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 136

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     137 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVFKTYN 196

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
           |||||||| | ||||||||||||||||| |||||||:||||||||||||||||||||||||
Db     197 VAMDYPTLFLAVWNFGAVGMVCIHWKGPLVLQQAYLIVISALMALVFIKYLPEWSAWVIL 256

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     257 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 316

Qy     241 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 300
           |||||||||||||||||||| || |  |||||| ||||||||||||||||||||||||||
Db     317 LPYDPEMEEDSYDSFGEPSYPEAFEAPQPGYPGEEPEEEEERGVKLGLGDFIFYSVLVGK 376

Qy     301 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 360
           |||||:|||:||||||:||||||||||||||||||||||||||||||||||||||||||||
Db     377 AAATGNGDWSTTLACFIAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 436
```

```
Qy    121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
          :|||||| : :|||||||:|||||||| ||||||||||||||||||||||||||||||
Db    200 IAMDYPTLFMVIWNFGAVGMICIHWKGPLQLQQAYLIMISALMALVFIKYLPEWSAWVIL 259

Qy    181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGAL- 240
          ||||||||:|||||||||||||||||||||||||||||||:||||||    :: | :
Db    260 GAISVYDLLAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMMWTVGMAD-SATADGRMN 318

Qy    240 -QLPY----DPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFY 294
           |: :    ||    : :   |           |       : :|||||||||||||||
Db    319 QQVQHIDRNTPEGANSTVEDAAETRIQ-------TQSNLSSEDPDEERGVKLGLGDFIFY 371

Qy    295 SVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFST 354
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    372 SVLVGKAAATASGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFST 431

Qy    355 DNLVRPFMDTLASHQLYI 372
          ||:|||||||||||||:||
Db    432 DNIVRPFMDTLASHQMYI 449


RESULT   6
PSN1_MOUSE
ID   PSN1_MOUSE     STANDARD;      PRT;   467 AA.
AC   P49769;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-DEC-1999 (Rel. 39, Last annotation update)
DE   PRESENILIN 1 (PS-1) (S182 PROTEIN).
GN   PSEN1 OR PSNL1 OR AD3H.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 95319502.
RA   SHERRINGTON R., ROGAEV E.I., LIANG Y., ROGAEVA E.A., LEVESQUE G.,
RA   IKEDA M., CHI H., LIN C., LI G., HOLMAN K., TSUDA T., MAR L.,
RA   FONCIN J.-F., BRUNI A.C., MONTESI M.P., SORBI S., RAINERO I.,
RA   PINESSI L., NEE L., CHUMAKOV I., POLLEN D., BROOKES A.,
RA   SANSEAU P., POLINSKY R.J., WASCO W., DA SILVA H.A.R., HAINES J.L.,
RA   PERICAK-VANCE M.A., TANZI R.E., ROSES A.D., FRASER P.E.,
RA   ROMMENS J.M., ST GEORGE-HYSLOP P.H.;
RT   "Cloning of a gene bearing missense mutations in early-onset familial
RT   Alzheimer's disease.";
RL   Nature 375:754-760(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SVJ;
RX   MEDLINE; 97442406.
RA   MITSUDA N., ROSES A.D., VITEK M.P.;
RT   "Transcriptional regulation of the mouse presenilin-1 gene.";
RL   J. Biol. Chem. 272:23489-23497(1997).
CC   -!- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE
CC       EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE.
CC       MAY FUNCTION IN THE CYTOPLASMIC PARTITIONING OF PROTEINS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L42177; AAC42094.1; -.
DR   EMBL; AF007560; AAB72049.1; -.
DR   MGD; MGI:1202717; PSEN1.
DR   PFAM; PF01080; Presenilin; 1.
KW   Transmembrane; Glycoprotein.
FT   TRANSMEM     83    103       POTENTIAL.
FT   TRANSMEM    133    153       POTENTIAL.
FT   TRANSMEM    161    181       POTENTIAL.
FT   TRANSMEM    195    215       POTENTIAL.
FT   TRANSMEM    221    241       POTENTIAL.
FT   TRANSMEM    244    264       POTENTIAL.
FT   TRANSMEM    281    301       POTENTIAL.
FT   TRANSMEM    408    428       POTENTIAL.
FT   TRANSMEM    433    453       POTENTIAL.
FT   CARBOHYD     30     30       POTENTIAL.
FT   CARBOHYD    279    279       POTENTIAL.
FT   CARBOHYD    405    405       POTENTIAL.
SQ   SEQUENCE   467 AA;  52639 MW;  3AE1350D CRC32;


  Query Match             74.4%;  Score 1431.5;  DB 1;  Length 467;
  Best Local Similarity   71.2%;  Pred. No. 4.2e-91;
  Matches  284;  Conservative   36;  Mismatches   50;  Indels   29;  Gaps    4;

Qy      1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
          |||||||||||||||||||||:||||||||| ||| |:|||||||||||| :|||| |
Db     71 EELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDTETVGQRAL 130

Qy     61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
          :|:|| ||||||:||| |||||||||| || |||:|||:||| |::|||||| ||||
Db    131 HSILNAAIMISVIVIMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYLGEVFKTYN 190

Qy    121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
          ||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||||||:||:||
Db    191 VAVDYVTVALLIWNFGVVGMIAIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLIL 250

Qy    181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240
            ||||||||||||||||||||||||| :|||||||| ||| | ||: || :|   :
Db    251 AVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGDPEAQ--RR 308

Qy    241 LPYDPEME--------EDSYDSFGEPSYPEVFEPPLTGYPG------------EEL---- 277
          :| :|:          :||     :  : | :|     : |             :||
Db    309 VPKNPKYNTQRAERETQDSGSGNDDGGFSEEWEAQRDSHLGPHRSTPESRAAVQELSGSI 368

Qy    277 ---EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVF 333
             |: |||||||||||||||||||:|| |||||||:|||||||||||||||||||:|
Db    369 LTSEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIF 428

Qy    334 KKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
          |||||||||||||:|||:|| ||:|||| || || ||
Db    429 KKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467


RESULT   7
PSN1_XENLA
ID   PSN1_XENLA     STANDARD;      PRT;   433 AA.
AC   O12976;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   PRESENILIN ALPHA.
GN   PS-ALPHA.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 97223465.
RA   TSUJIMURA A., YASOJIMA K., HASHIMOTO-GOTOH T.;
RT   "Cloning of Xenopus presenilin-alpha and -beta cDNAs and their
RT   differential expression in oogenesis and embryogenesis.";
RL   Biochem. Biophys. Res. Commun. 231:392-396(1997).
CC   -!- FUNCTION: MAY PLAY A ROLE IN NEGATIVE REGULATION OF APOPTOTIC
CC       CASCADES DURING OOGENESIS AND EMBRYOGENESIS, AND IN
CC       DEVELOPMENTALLY MATURED TISSUES SUCH AS BRAIN TISSUE.
```

```
Db    191 VAVDYITVALLIWNFGVVGMIAIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLIL 250

Qy    181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240
          ||||||||||||||||||||||||||| :|||||||| ||| | ||: || :|   :
Db    251 AVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGDPEAQ--RR 308

Qy    241 LPYDP---------EMEEDSYDSFGEPSYPEVFEPPLTGYPG------------EEL--- 277
          :| :|         |  :|:     :  : | :|     : |            :||
Db    309 VPKNPKYSTQGTEREETQDTGTGSDDGGFSEEWEAQRDSHLGPHRSTPESRAAVQELSGS 368

Qy    277 ----EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAV 332
              |: ||||||||||||||||||||:|| |||||||:|||||||||||||||||||:
Db    369 ILTSEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAI 428

Qy    333 FKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
          ||||||||||||||||||:|| ||:|||| || || ||
Db    429 FKKALPALPISITFGLIFYFATDYLVQPFMDQLAFHQFYI 468


RESULT   9
PSN1_HUMAN
ID   PSN1_HUMAN     STANDARD;      PRT;   467 AA.
AC   P49768; Q14762; Q15719; Q15720;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   PRESENILIN 1 (PS-1) (S182 PROTEIN).
GN   PSEN1 OR PSNL1 OR AD3 OR PS1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.; AND VARIANTS AD (FORMS I-467 AND I-463).
RC   TISSUE=BRAIN;
RX   MEDLINE; 95319502.
RA   SHERRINGTON R., ROGAEV E.I., LIANG Y., ROGAEVA E.A., LEVESQUE G.,
RA   IKEDA M., CHI H., LIN C., LI G., HOLMAN K., TSUDA T., MAR L.,
RA   FONCIN J.-F., BRUNI A.C., MONTESI M.P., SORBI S., RAINERO I.,
RA   PINESSI L., NEE L., CHUMAKOV I., POLLEN D., BROOKES A.,
RA   SANSEAU P., POLINSKY R.J., WASCO W., DA SILVA H.A.R., HAINES J.L.,
RA   PERICAK-VANCE M.A., TANZI R.E., ROSES A.D., FRASER P.E.,
RA   ROMMENS J.M., ST GEORGE-HYSLOP P.H.;
RT   "Cloning of a gene bearing missense mutations in early-onset familial
RT   Alzheimer's disease.";
RL   Nature 375:754-760(1995).
RN   [2]
RP   SEQUENCE FROM N.A. (FORMS I-463 AND I-374).
RC   TISSUE=BLOOD, AND BRAIN;
RX   MEDLINE; 96193901.
RA   SAHARA N., YAHAGI Y.-I., TAKAGI H., KONDO T., OKOCHI M., USAMI M.,
RA   SHIRASAWA T., MORI H.;
RT   "Identification and characterization of presenilin I-467, I-463 and
RT   I-374.";
RL   FEBS Lett. 381:7-11(1996).
RN   [3]
RP   SEQUENCE OF 1-113 FROM N.A.
RA   TSUJIMURA A., HASHIMOTO-GOTOH T.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   REVIEW ON VARIANTS.
RX   MEDLINE; 98180715.
RA   CRUTS M., VAN BROECKHOVEN C.;
RT   "Presenilin mutations in Alzheimer's disease.";
RL   Hum. Mutat. 11:183-190(1998).
RN   [5]
RP   VARIANTS AD THR-143 AND ALA-384.
RX   MEDLINE; 96177673.
RA   CRUTS M., BACKHOVENS H., WANG S.-Y., VAN GASSEN G., THEUNS J.,
RA   DE JONGHE C., WEHNERT A., DE VOECHT J., DE WINTER G., CRAS P.,
RA   BRUYLAND M., DATSON N., WEISSENBACH J., DEN DUNNEN J.T., MARTIN J.-J.,
RA   HENDRIKS L., VAN BROECKHOVEN C.;
RT   "Molecular genetic analysis of familial early-onset Alzheimer's
RT   disease linked to chromosome 14q24.3.";
RL   Hum. Mol. Genet. 4:2363-2372(1995).
RN   [6]
RP   VARIANTS AD L-82; H-115; T-139; R-163; T-231; L-264; V-392 AND Y-410.
RX   MEDLINE; 96177674.
RA   CAMPION D., FLAMAN J.-M., BRICE A., HANNEQUIN D., DUBOIS B.,
RA   MARTIN C., MOREAU V., CHARBONNIER F., DIDIERJEAN O., TARDIEU S.,
RA   PENET C., PUEL M., PASQUIER F., LE DOZE F., BELLIS G., CALENDA A.,
RA   HEILIG R., MARTINEZ M., MALLET J., BELLIS M., CLERGET-DARPOUX F.,
RA   AGID Y., FREBOURG T.;
RT   "Mutations of the presenilin I gene in families with early-onset
RT   Alzheimer's disease.";
RL   Hum. Mol. Genet. 4:2373-2377(1995).
RN   [7]
RP   VARIANTS AD VAL-260; VAL-285 AND VAL-392.
RX   MEDLINE; 95379971.
RA   ROGAEV E.I., SHERRINGTON R., ROGAEVA E.A., LEVESQUE G., IKEDA M.,
RA   LIANG Y., CHI H., LIN C., HOLMAN K., TSUDA T., MAR L., SORBI S.,
RA   NACMIAS B., PIACENTINI S., AMADUCCI L., CHUMAKOV I., COHEN D.,
RA   LANNFELT L., FRASER P.E., ROMMENS J.M., ST GEORGE-HYSLOP P.H.;
RT   "Familial Alzheimer's disease in kindreds with missense mutations in
RT   a gene on chromosome 1 related to the Alzheimer's disease type 3
RT   gene.";
RL   Nature 376:775-778(1995).
RN   [8]
RP   VARIANTS AD V-139; V-146; Y-163; T-267; A-280 AND G-280.
RX   MEDLINE; 96024664.
RA   CLARK R.F., HUTTON M., FULDNER R.A., FROELICH S., KARRAN E.,
RA   TALBOT C., CROOK R., LENDON C., PRIHAR G., HE C., KORENBLAT K.,
RA   MARTINEZ A., WRAGG M., BUSFIELD F., BEHRENS M.I., MYERS A., NORTON J.,
RA   MORRIS J., MEHTA N., PEARSON C., LINCOLN S., BAKER M., DUFF K.,
RA   ZEHR C., PEREZ-TUR J., HOULDEN H., RUIZ A., OSSA J., LOPERA F.,
RA   ARCOS M., MADRIGAL L., COLLINGE J., HUMPHREYS C., ASWORTH T.,
RA   SARNER S., FOX N., HARVEY R., KENNEDY A., ROQUES P., CLINE R.T.,
RA   PHILLIPS C.A., VENTER J.C., FORSEL L., AXELMAN K., LILIUS L.,
RA   JOHNSTON J., COWBURN R., VIITANEN M., WINBLAD B., KOSIK K., HALTIA M.,
RA   POYHONEN M., DICKSON D., MANN D., NEARY D., SNOWDEN J., LANTOS P.,
RA   LANNFELT L., ROSSOR M., ROBERTS G.W., ADAMS M.D., HARDY J., GOATE A.;
RT   "The structure of the presenilin 1 (S182) gene and identification of
RT   six novel mutations in early onset AD families. Alzheimer's Disease
RT   Collaborative Group.";
RL   Nat. Genet. 11:219-222(1995).
RN   [9]
RP   VARIANTS AD PHE-96; ARG-163 AND THR-213.
RX   MEDLINE; 96310408.
RA   KAMINO K., SATO S., SAKAKI Y., YOSHIIWA A., NISHIWAKI Y., TAKEDA H.,
RA   TANABE H., NISHIMURA T., LI K., ST GEORGE-HYSLOP P.H., MIKI T.,
RA   OGIHARA T.;
RT   "Three different mutations of presenilin 1 gene in early-onset
RT   Alzheimer's disease families.";
RL   Neurosci. Lett. 208:195-198(1996).
RN   [10]
RP   VARIANT AD ALA-280.
RX   MEDLINE; 97442268.
RA   LENDON C.L., MARTINEZ A., BEHRENS I.M., KOSIK K.S., MADRIGAL L.,
RA   NORTON J., NEUMAN R., MYERS A., BUSFIELD F., WRAGG M., ARCOS M.,
RA   ARANGO VIANA J.C., OSSA J., RUIZ A., GOATE A.M., LOPERA F.;
RT   "E280A PS-1 mutation causes Alzheimer's disease but age of onset is
RT   not modified by ApoE alleles.";
RL   Hum. Mutat. 10:186-195(1997).
RN   [11]
RP   VARIANTS AD THR-233 AND THR-278.
RX   MEDLINE; 97316242.
RA   KWOK J.B.J., TADDEI K., HALLUPP M., FISHER C., BROOKS W.S., BROE G.A.,
RA   HARDY J., FULHAM M.J., NICHOLSON G.A., STELL R.,
RA   ST GEORGE-HYSLOP P.H., FRASER P.E., KAKULAS B., CLARNETTE R.,
RA   RELKIN N., GANDY S.E., SCHOFIELD P.R., MARTINS R.N.;
RT   "Two novel (M233T and R278T) presenilin-1 mutations in early-onset
RT   Alzheimer's disease pedigrees and preliminary evidence for
RT   association of presenilin-1 mutations with a novel phenotype.";
RL   NeuroReport 8:1537-1542(1997).
RN   [12]
RP   VARIANT GLY-318.
RX   MEDLINE; 99115106.
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z71333; CAA95930.1; -.
DR   PFAM; PF01080; Presenilin; 1.
KW   Transmembrane; Alternative splicing; Glycoprotein.
FT   TRANSMEM     83    103       POTENTIAL.
FT   TRANSMEM    133    153       POTENTIAL.
FT   TRANSMEM    161    181       POTENTIAL.
FT   TRANSMEM    191    211       POTENTIAL.
FT   TRANSMEM    221    241       POTENTIAL.
FT   TRANSMEM    244    264       POTENTIAL.
FT   TRANSMEM    281    301       POTENTIAL.
FT   TRANSMEM    408    428       POTENTIAL.
FT   TRANSMEM    433    453       POTENTIAL.
FT   CARBOHYD    279    279       POTENTIAL.
FT   CARBOHYD    405    405       POTENTIAL.
FT   VARSPLIC     26     29       MISSING (IN ISOFORM I-463).
SQ   SEQUENCE   467 AA;  52384 MW;  A841A0B7 CRC32;


  Query Match           73.5%;  Score 1413.5;  DB 1;  Length 467;
  Best Local Similarity   71.0%;  Pred. No. 7.2e-90;
  Matches  282;  Conservative   34;  Mismatches   56;  Indels   25;  Gaps    4;

Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           ||||||||||||||||||||||:|||||||||| ||| |:|||||||||||| :|||| |
Db      71 EELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDTETVGQRAL 130

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           :||||  |||||||||||| ||||||||||| || |||:|||:||| |::|||||| ||||
Db     131 HSVLNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYLGEVFKTYN 190

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
           ||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||||||:||:||
Db     191 VAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLIL 250

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGAL- 240
             |||||||||||||||||||||||||| :|||||||| ||| | ||: || :|  :
Db     251 AVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVS 310

Qy     240 -QLPYD------------PEMEEDSYDSFGEPSYPEVFEP----PLTGYPGEEL------ 277
              |:             || ::  :    |        |     ::    :|:
Db     311 KNTKYNAQGTEREAQASVPENDDGGFSEEWEAQRDSQLGPHRSTSVSRAAVQEISSSIPA 370

Qy     277 -EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKK 335
            |: ||||||||||||:||||||||:|| |||||||:||||||||||||||||:|||
Db     371 SEDPEERGVKLGLGDFVFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKK 430

Qy     336 ALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
           |||||||||||||:|||:|| ||:|||| || || ||
Db     431 ALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467


RESULT  11
PSN_DROME
ID   PSN_DROME      STANDARD;      PRT;   541 AA.
AC   O02194; O02395; O76802;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   PRESENILIN HOMOLOG (DPS) (DMPS).
GN   PS.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE FROM N.A. (LONG ISOFORM).
RX   MEDLINE; 97285868.
RA   BOULIANNE G.L., LIVNE-BAR I., HUMPHREYS J.M., LIANG Y., LIN C.,
RA   ROGAEV E., ST GEORGE-HYSLOP P.;
RT   "Cloning and characterization of the Drosophila presenilin
RT   homologue.";
RL   NeuroReport 8:1025-1029(1997).
RN   [2]
RP   SEQUENCE FROM N.A. (SHORT ISOFORM).
RC   STRAIN=CANTON-S;
RA   HONG C.S., KOO E.H.;
RT   "Isolation and characterization of Drosophila presenilin homolog.";
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
RA   YE Y., FORTINI M.E.;
RT   "Characterization of Drosophila presenilin and its colocalization
RT   with Notch during development.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U77934; AAB61139.1; -.
DR   EMBL; U78084; AAB53369.1; -.
DR   EMBL; AF084184; AAC33129.1; -.
DR   EMBL; AF084184; AAC33128.1; -.
DR   FLYBASE; FBgn0019947; PS.
DR   PFAM; PF01080; Presenilin; 1.
KW   Transmembrane; Glycoprotein; Alternative splicing.
FT   TRANSMEM    107    127       POTENTIAL.
FT   TRANSMEM    155    175       POTENTIAL.
FT   TRANSMEM    183    203       POTENTIAL.
FT   TRANSMEM    217    237       POTENTIAL.
FT   TRANSMEM    243    263       POTENTIAL.
FT   TRANSMEM    266    286       POTENTIAL.
FT   TRANSMEM    304    324       POTENTIAL.
FT   TRANSMEM    482    502       POTENTIAL.
FT   TRANSMEM    507    527       POTENTIAL.
FT   CARBOHYD    129    129       POTENTIAL.
FT   CARBOHYD    339    339       POTENTIAL.
FT   CARBOHYD    410    410       POTENTIAL.
FT   VARSPLIC    384    397       MISSING (IN SHORT ISOFORM).
FT   CONFLICT     80     81       GG -> RS (IN REF. 2).
SQ   SEQUENCE   541 AA;  59304 MW;  796C4FE0 CRC32;


  Query Match           58.6%;  Score 1127.5;  DB 1;  Length 541;
  Best Local Similarity   50.8%;  Pred. No. 2.9e-70;
  Matches  231;  Conservative   61;  Mismatches   74;  Indels   89;  Gaps    7;

Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           ||  |||||:||| |||||:|||:|||||| |: ||    :  |:|||| | :|     :
Db      93 EEQGLKYGAQHVIKLFVPVSLCMLVVVATINSISFYNSTDVYLLYTPFHEQSPEPSVKFW 152

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           ::: |:||::||:||||  |:|||| ||||: ||||||:|| ||||:|||:|| |:|: ||
Db     153 SALANSLILMSVVVVMTFLLIVLYKKRCYRIIHGWLILSSFMLLFIFTYLYLEELLRAYN 212

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
           : ||||| || :|||| |||: |||:||| ||| ||| ::|||||||||||||:|| :|
Db     213 IPMDYPTALLIMWNFGVVGMMSIHWQGPLRLQQGYLIFVAALMALVFIKYLPEWTAWAVL 272

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTV----------GMAK 230
             |||::||:||| |:||||:|||||||||| |||||||| :|: :           |
Db     273 AAISIWDLIAVLSPRGPLRILVETAQERNEQIFPALIYSSTVVYALVNTVTPQQSQATAS 332

Qy     231 LDPSS----------QGALQLP-------------YDPEMEEDSYDSFGEPSYPEV-FEP 266
             |||          | :|  |             :  : : |          | | |:
Db     333 SSPSSSNSTTTTRATQNSLASPEAAAASGQRTGNSHPRQNQRDDGSVLATEGMPLVTFKS 392
```

```
  Query Match           50.5%;  Score 972;  DB 1;  Length 461;
  Best Local Similarity  52.8%;  Pred. No. 1e-59;
  Matches  211;  Conservative   53;  Mismatches   74;  Indels   62;  Gaps    9;

Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQ-LIYTPFTEDTPSVGQRL 59
           ||  ||||| ||| ||||||:||| :|| |: :: ||:: ||: |:||||  :| |: ::
Db      39 EEAELKYGASHVIHLFVPVSLCMALVVFTMNTITFYSQNNGRHLLYTPFVRETDSIVEKG 98

Qy      60 LNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTY 119
           | |: | |:|: |:|:||: |:| ||||: || ||||||:|| :|||||| ||: ||||::
Db      99 LMSLGNALVMLCVVVLMTVLLIVFYKYKFYKLIHGWLIVSSFLLLFLFTTIYVQEVLKSF 158

Qy     120 NVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVI 179
           :|:     :|  : |:| :||:|||||||| ||| ||| :|||||||||||||||: | :
Db     159 DVSPSALLVLFGLGNYGVLGMMCIHWKGPLRLQQFYLITMSALMALVFIKYLPEWTVWFV 218

Qy     180 LGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGAL 239
           |  |||:|||||| |||||| |||||||||||||||||||||| :::   :         |:
Db     219 LFVISVWDLVAVLTPKGPLRYLVETAQERNEPIFPALIYSSGVIYPYVLVT-------AV 271

Qy     240 QLPYDPEMEEDSYDSF------GE-------PSYPEVFEPP------------LTGYPGE 274
           :   ||  |  | ||       ||       |  |:|  |             |   |
Db     272 ENTTDPR-EPTSSDSNTSTAFPGEASCSSETPKRPKVKRIPQKVQIESNTTASTTQNSGV 330

Qy     275 ELE-----------------EEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACF 316
            :|                  ||||||||||||||||||:|||::    |||||:||:
Db     331 RVERELAAERPTVQDANFHRHEEEERGVKLGLGDFIFYSVLLGKASS--YFDWNTTIACY 388

Qy     317 VAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDN 356
           |||||||| ||:||||||:|||||         | || |:
Db     389 VAILIGLCFTLVLLAVFKRALPALQ--------FPFSPDS 420


RESULT  14
HOP1_CAEEL
ID   HOP1_CAEEL     STANDARD;      PRT;   358 AA.
AC   O02100;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-DEC-1999 (Rel. 39, Last annotation update)
DE   INTEGRAL MEMBRANE PROTEIN HOP-1.
GN   HOP-1 OR C18E3.8.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC   Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE; 98004548.
RA   LI X., GREENWALD I.;
RT   "HOP-1, a Caenorhabditis elegans presenilin, appears to be
RT   functionally redundant with SEL-12 presenilin and to facilitate LIN-12
RT   and GLP-1 signaling.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:12204-12209(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   CONNELL M., MAGGI L.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAY FACILITATE LIN-12 AND GLP-1 MEDIATED RECEPTION OF
CC       INTERCELLULAR SIGNALS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF021905; AAB84394.1; -.
DR   EMBL; AF000265; AAB52948.1; -.
DR   WORMPEP; C18E3.8; CE08317.
DR   PFAM; PF01080; Presenilin; 1.
KW   Transmembrane.
FT   TRANSMEM     13     33       POTENTIAL.
FT   TRANSMEM     58     78       POTENTIAL.
FT   TRANSMEM     87    107       POTENTIAL.
FT   TRANSMEM    116    136       POTENTIAL.
FT   TRANSMEM    149    169       POTENTIAL.
FT   TRANSMEM    171    191       POTENTIAL.
FT   TRANSMEM    298    318       POTENTIAL.
FT   TRANSMEM    322    342       POTENTIAL.
SQ   SEQUENCE   358 AA;  39864 MW;  FC6305ED CRC32;


  Query Match           27.3%;  Score 524.5;  DB 1;  Length 358;
  Best Local Similarity  30.5%;  Pred. No. 3e-29;
  Matches  117;  Conservative   81;  Mismatches  136;  Indels   49;  Gaps    8;

Qy       7 YGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPF----TEDTPSVGQRLLNS 62
           |  | :  :  || :||: |   :|  :   ::  :::|  |    | |: ::
Db       8 YSGKTITGVLYPVAICMLFVAINVKLSQPEQQEQSKVVYGLFHSYDTADSGTI------- 61

Qy      63 VLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYNVA 122
             || :|  ::: |   |  |: : || |  ::: :|: :| :::  :   : :  ::
Db      61 ---TLYLIGFLILTTSLGVFCYQMKFYKAIKVYVLANSIGILLVYSVFHFQRIAEAQSIP 117

Qy     123 MDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVILGA 182
           :  ||    :  || :|: |:|||    | | ||||:: | |:  :  ||:|: |: | |
Db     118 VSVPTFFFLILQFGGLGITCLHWKSHRRLHQFYLIMLAGLTAIFILNILPDWTVWMALTA 177

Qy     183 ISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVW-----------TVGMAK 230
           || :|:|||| | |||:||||||  | :  |||::|:|:                :  : :
Db     178 ISFWDIVAVLTPCGPLKMLVETANRRGDDKFPAILYNSSSYVNEVDSPDTTRSNSTPLTE 237

Qy     231 LDPSSQGALQLPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERG-VKLGLG 289
           : ||   |       :| ||           :  ||:   | :  |  |  | ::||:|
Db     238 FNNSSSSRL-------LESDS----------LLRPPV--IPRQIREVREVEGTIRLGMG 277

Qy     290 DFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLI 349
           ||:|||:::|    |       | :||||: |:|| :|| :: : : ||||||  :    |
Db     278 DFVFYSLMLGNTVQT--CPLPTVVACFVSNLVGLTITLPIVTLSQTALPALPFPLAIAAI 335

Qy     350 FYFSTDNLVRPFMDTLASHQLYI 372
           ||||:   : || |   |  : |
Db     336 FYFSSHIALTPFTDLCTSQLILI 358


RESULT  15
PSNH_ARATH
ID   PSNH_ARATH     STANDARD;      PRT;   453 AA.
AC   O64668;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-DEC-1999 (Rel. 39, Last annotation update)
DE   PRESENILIN HOMOLOG.
GN   F22O13.19.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC   core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC   Arabidopsis.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   SHINN P., BUEHLER E., DEWAR K., FENG J., KIM C., LI Y., SUN H.,
RA   CONWAY A.B., CONWAY A.R., KURTZ D., OJI O., SHEN Y.K., TORIUMI M.,
RA   VYSOTSKAIA V., YU G., DAVIS R.W., FEDERSPIEL N.A., THEOLOGIS A.,
RA   ECKER J.R.;
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 18, 2000, 22:07:43 ; Search time 49.26 Seconds
(without alignments)
523.595 Million cell updates/sec

Title: US-08-509-359B-138
Perfect score: 1923
Sequence: 1 EELTLKYGAKHVIMLFVPVT.....STDNLVRPFMDTLASHQLYI 372

Scoring table: BLOSUM62

Searched: 225878 seqs, 69334122 residues

Database : SPTREMBL_12:*

Word size : 0

Number of hits that pass the threshold : 225878
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1896	98.6	449	6	Q9XT96	Q9xt96 bos taurus
2	1411.5	73.4	478	6	Q9XT97	Q9xt97 bos taurus
3	1398	72.7	456	13	Q9W6T7	Q9w6t7 brachydanio
4	1355	70.5	384	13	O73869	O73869 cyprinus ca
5	576.5	30.0	272	5	O96340	O96340 drosophila
6	402.5	20.9	184	4	O95465	O95465 homo sapien
7	113.5	5.9	406	5	Q19737	Q19737 caenorhabdi
8	112	5.8	4578	13	O42181	O42181 fugu rubrip
9	110.5	5.7	320	8	Q34086	Q34086 coccyzus er
10	110.5	5.7	381	8	Q35425	Q35425 phascolosor
11	108.5	5.6	380	8	Q9ZZC9	Q9zzc9 upupa epops
12	107.5	5.6	748	2	Q9Z577	Q9z577 streptomyce
13	107	5.6	381	8	Q33723	Q33723 antechinus
14	106.5	5.5	382	8	Q34340	Q34340 didelphis m
15	106.5	5.5	318	11	O35294	O35294 rattus norv
16	106.5	5.5	303	11	P97829	P97829 rattus norv
17	106	5.5	652	5	Q93346	Q93346 caenorhabdi
18	106	5.5	382	8	Q34279	Q34279 didelphis a
19	106	5.5	382	8	Q34677	Q34677 glironia ve
20	105	5.5	444	2	Q9X2N3	Q9x2n3 arthrobacte
21	105	5.5	379	8	Q34428	Q34428 echimys did
22	104	5.4	382	8	Q35561	Q35561 philander o
23	104	5.4	379	8	Q36096	Q36096 trinomys pa
24	104	5.4	379	8	Q34430	Q34430 echimys did
25	104	5.4	826	10	O80739	O80739 arabidopsis
26	103	5.4	266	8	Q34429	Q34429 echimys did
27	103	5.4	266	8	Q37099	Q37099 echimys did
28	103	5.4	380	8	Q9XNM7	Q9xnm7 microtus lo
29	102.5	5.3	382	8	Q34973	Q34973 micoureus d
30	102.5	5.3	382	8	Q35075	Q35075 marmosa rub
31	102.5	5.3	382	8	Q35076	Q35076 micoureus r
32	102	5.3	444	2	O85080	O85080 arthrobacte
33	102	5.3	266	8	Q34425	Q34425 echimys did
34	102	5.3	266	8	Q34427	Q34427 echimys did
35	102	5.3	379	8	Q34424	Q34424 echimys did
36	102	5.3	379	8	Q33940	Q33940 akodon uric
37	102	5.3	381	8	Q9XP75	Q9xp75 sminthopsis
38	101.5	5.3	1681	11	Q62467	Q62467 mus musculu
39	101	5.3	381	8	Q9XP88	Q9xp88 sminthopsis
40	101	5.3	380	8	Q9XNL9	Q9xnl9 ondatra zib
41	100.5	5.2	382	8	Q35518	Q35518 philander m
42	100.5	5.2	324	11	Q61735	Q61735 mus musculu
43	100	5.2	380	8	Q9XKB9	Q9xkb9 phaethon le
44	99.5	5.2	379	5	Q93127	Q93127 balanus amp
45	99.5	5.2	382	8	Q35558	Q35558 philander o

ALIGNMENTS

RESULT 1
Q9XT96
ID Q9XT96 PRELIMINARY; PRT; 449 AA.
AC Q9XT96;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE PRESENILIN 2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae;
OC Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA SAHARA N., SHIRASAWA T., MORI H.;
RT "Molecular cloning of bovine presenilin 2 gene.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF038937; AAD39024.1; -.
SQ SEQUENCE 449 AA; 50301 MW; A3DA878F CRC32;

Query Match 98.6%; Score 1896; DB 6; Length 449;
Best Local Similarity 98.4%; Pred. No. 9.9e-131;
Matches 366; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
           ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db      78 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFSEDTPSVGQRLL 137

Qy      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120
           ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     138 NSVLNTLIMISVIVTMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 197

Qy     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     198 VAMDYPTLFLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 257

Qy     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     258 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 317

Qy     241 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 300
           |||||||||||||||||||:|||||| ||||||||||||||||||||||||||||||||||
Db     318 LPYDPEMEEDSYDSFGEPSYPDVFEPPLPGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 377

Qy     301 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 360
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1 IMLFIPVTLCMVVVVATIKSVSFYTQKDGQQLIYTPFREDTETVGQRALNSMLNAAIMIS 60

Qy     72 VIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYNVAMDYPTLLLT 131
          |||||||: |||||||||||| | |||  |:|:||| |::||||||| ||||||||| || :
Db     61 VIVVMTLVLVVLYKYRCYKVIQGWLFFSNLLLLFFFSFIYLGEVFKTYNVAMDYFTLAVI 120

Qy    132 VWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVILGAISVYDLVAV 191
          :|||| |||:|||||||| |||||||||||||||||||||||||:||:|| |||||||:||
Db    121 IWNFGVVGMICIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLILAAISVYDLLAV 180

Qy    192 LCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMA-KLDPSSQGALQLP-------- 243
          ||||||||:|||||||||||||||||||| |||   ||   :| :  :  :|
Db    181 LCPKGPLRILVETAQERNEPIFPALIYSSTMVWLFNMADSAEPRNNSSHPVPQQENQDAV 240

Qy    243 -----------YDPEMEEDSYDSFG-----EPSYPEVFEPPLTGYPGEELEEEEERGVKL 286
                    : |          |     | |  |: | |    |   :|::|||||||
Db    241 APTAQPEDDGGFTPAWVNQQQHQLGPMQSTEDSRREIQELPSARPP--PVEDDEERGVKL 298

Qy    287 GLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITF 346
          |||||||||:|||||:|| |||||||:||||||||||||||||:|||||||||||||||
Db    299 GLGDFIFYSMLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITF 358

Qy    347 GLIFYFSTDNLVRPFMDTLASHQLYI 372
          ||:|||:|||||||||| || || ||
Db    359 GLVFYFATDNLVRPFMDQLAVHQFYI 384


RESULT   5
O96340
ID   O96340      PRELIMINARY;      PRT;   272 AA.
AC   O96340;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   PRESENILIN (FRAGMENT).
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CANTON S;
RX   MEDLINE; 98331525.
RA   MARFANY G., DEL-FAVERO J., VALERO R., DE JONGHE C., WOODROW S.,
RA   HENDRIKS L., VAN BROECKHOVEN C., GONZALEZ-DUARTE R.;
RT   "Identification of a Drosophila presenilin homologue: evidence of
RT   alternatively spliced forms.";
RL   J. Neurogenet. 12:41-54(1998).
DR   EMBL; AF017025; AAD01611.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   272 AA;  29456 MW;  606B9A5C CRC32;


  Query Match             30.0%;  Score 576.5;  DB 5;  Length 272;
  Best Local Similarity   47.8%;  Pred. No. 6.8e-35;
  Matches  133;  Conservative   29;  Mismatches   41;  Indels   75;  Gaps    8;

Qy    164 ALVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMV 223
          ||||||||||:|| :| |||::||:||| |:||||:|||||||||| |||||||| :|
Db      1 ALVFIKYLPEWTAWAVLAAISIWDLIAVLSPRGPLRILVETAQERNEQIFPALIYSSTVV 60

Qy    224 WTV----------GMAKLDPSS----------QGALQLPYDPEMEEDSYDSFGEPSYPE- 263
          : :          |   |||          | :|     ||    |    |  |:|
Db     61 YALVNTVTPQQSQATASSSPSSSNSTTTTRATQNSLA---SPEAAAASGQRTGN-SHPRQ 116

Qy    263 ---------VFEPPLTGYPGE---ELEE------------------------------ 279
                  |    |:  |    | |
Db    117 NQRDDGSVLATEAEAAGFTQEWSANLSERVARRQIEVQSTQSGNAQRSNEYRTVTAPDQN 176

Qy    279 ----EEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFK 334
              :||||:|||||||||||||||::   ||| ||:|||||||||||||||:::
Db    177 HPDGQEERGIKLGLGDFIFYSVLVGKASS--YGDWTTTIACFVAILIGLCLTLLLLAIWR 234

Qy    335 KALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 372
          |||||||||||||||| |:|  :|:|||: |:: |::|
Db    235 KALPALPISITFGLIFCFATSAVVKPFMEDLSAKQVFI 272


RESULT   6
O95465
ID   O95465      PRELIMINARY;      PRT;   184 AA.
AC   O95465;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   MINILIN.
GN   PSN1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   POWELL C.S., GEGG M.E., PALMER M.S.;
RT   "Human presenilin 1 gene encodes an alternative protein-minilin.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ008005; CAA07825.1; -.
SQ   SEQUENCE   184 AA;  21073 MW;  5C6FBAEE CRC32;


  Query Match             20.9%;  Score 402.5;  DB 4;  Length 184;
  Best Local Similarity   73.5%;  Pred. No. 2.1e-22;
  Matches   83;  Conservative    9;  Mismatches   16;  Indels    5;  Gaps    1;

Qy      1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60
          |||||||||||||||||||||:||||||||| ||| |:|||||||||||| :|||| |
Db     71 EELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDTETVGQRAL 130

Qy     61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLG 113
          :|:||  |||||||||| ||||||||||     : |    ||    :::||
Db    131 HSILNAAIMISVIVVMTILLVVLYKYRCYK-----VSMRHRSLLSTLFFLWLG 178


RESULT   7
Q19737
ID   Q19737      PRELIMINARY;      PRT;   406 AA.
AC   Q19737; Q22692;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   F22E10.5 PROTEIN.
GN   F22E10.5.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC   Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   GARDNER A.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z67882; CAA91804.1; -.
DR   EMBL; Z50797; CAA91804.1; JOINED.
DR   EMBL; Z50797; CAA90677.1; -.
DR   EMBL; Z67882; CAA90677.1; JOINED.
DR   PROSITE; PS00379; CDP_ALCOHOL_P_TRANSF; 1.
SQ   SEQUENCE   406 AA;  45628 MW;  07336492 CRC32;


  Query Match              5.9%;  Score 113.5;  DB 5;  Length 406;
  Best Local Similarity   24.7%;  Pred. No. 0.54;
  Matches   64;  Conservative   31;  Mismatches   71;  Indels   93;  Gaps   13;

Qy      7 YGAKHVIMLFVPVTLCM----------IVVVATIKSVRFY-----TEKNGQLIYTPF--T 49
          :|   :  :|| : :|           :::|: |  | ||     |   ||| :: |  |
Db    136 HGCDSMTQVFVTLNICYAMSLGTVPYGVLIVSVISVVMFYCAHWSTYCTGQLRFSKFDVT 195

Qy     50 EDTPSVGQRLL------NSVLNTLI----MISVIVVMTIFLVVLYKYRCYKFIHGWL--- 97
          |   ||   ||      | |  | |     :   || : ||| ||:      |||::
```

```
Qy    342 ISITFGLIFYFSTDNLVRPFMDTLASH 368
          :| |   :|:|  :||    :  | ||
Db    288 LSQT---LFWFLVENLF--ILKWLGSH 309


RESULT  10
Q35425
ID   Q35425      PRELIMINARY;      PRT;   381 AA.
AC   Q35425;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   CYTOCHROME B.
OS   Phascolosorex dorsalis.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Metatheria; Dasyuromorphia; Dasyuridae; Phascolosorex.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93096825.
RA   KRAJEWSKI C., DRISKELL A.C., BAVERSTOCK P.R., BRAUN M.J.;
RT   "Phylogenetic relationships of the thylacine (Mammalia: Thylacinidae)
RT   among dasyuroid marsupials: evidence from cytochrome b DNA
RT   sequences.";
RL   Proc. R. Soc. Lond., B, Biol. Sci. 250:19-27(1992).
CC   -!- CATALYTIC ACTIVITY: QH(2) + 2 FERRICYTOCHROME C = Q + 2
CC       FERROCYTOCHROME C.
CC   -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC       BOUND TO THE PROTEIN (BY SIMILARITY).
DR   EMBL; M99462; AAB61692.1; -.
DR   PFAM; PF00033; cytochrome_b_N; 1.
KW   Mitochondrion; Electron transport; Respiratory chain; Transmembrane;
KW   Heme.
SQ   SEQUENCE   381 AA;  42813 MW;  C6A2674F CRC32;


  Query Match             5.7%;  Score 110.5;  DB 8;  Length 381;
  Best Local Similarity   21.1%;  Pred. No. 0.84;
  Matches   81;  Conservative   54;  Mismatches  115;  Indels  133;  Gaps   18;

Qy     21 LCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFL 80
          :|:|: : |           |  :   :| ||                        :| |
Db     39 MCLIIQILT----------GLFLAMHYTSDT----------------------LTAFT 64

Qy     81 VVLYKYRCYKFIHGWLIM------SSLMLLFLFTYIYLGEVLKTYNVAMDYPTLLLTVWN 134
           |   : |    :|||:       :|:  : || :|  |    :|       |    ||
Db     65 SV--AHICRDVNYGWLLRNLHANGASMFFMCLFLHIGRGIYYGSY--------LYKETWN 114

Qy    135 FGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVILGAISVYDLVAVLCP 194
           |                  :|:: |:||  |: |:  |      ||  : :|:: :
Db    115 IG------------------VILLLAVMATAFVGYVLPWGQMSFWGATVITNLLSAIPY 155

Qy    195 KGPLRMLVE------------TAQERNEPIFPALIYSSAMVWTVGMAKL---DPS--SQ 236
           ||   | |              |       | | :: : |:|  : : :    :||  :
Db    156 IGP--TLAEWVWGGYAVDKATLTRFFAFHFILPFIVTALAIVHLLFLHETGSNNPSGINP 213

Qy    237 GALQLPYDPE-----------------MEEDSYDSFGEPSYPEVFEP--PLTGYPGEEL 276
           | ::|: |                    :   | || |:   |: | |  ||   |  :
Db    214 NADKIPFHPYYTIKDALGFMLLLSVLLLLTLFSPDSLGD---PDNFSPANPLNTPPHIKP 270

Qy    277 EEEEERGVKLGLGDFIF-YSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKK 335
           |            |:| |::|       |       ||   :||| | : ||  |  ::
Db    271 E----------WYFLFAYAILRSIPNKLGG-----VLALLASILILLIIPLLHTA-NQR 313

Qy    336 ALPALPISITFGLIFYFSTDNLV 358
          ::   ||| |   :|:  | ||:
Db    314 SMMFRPISQT---LFWILTANLI 333


RESULT  11
Q9ZZC9
ID   Q9ZZC9      PRELIMINARY;      PRT;   380 AA.
AC   Q9ZZC9;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   CYTOCHROME B.
GN   CYTB.
OS   Upupa epops.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Upupiformes; Upupidae; Upupa.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   ESPINOSA DE LOS MONTEROS A.;
RT   "Higher-level phylogeny of Trogoniformes.";
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U89189; AAD00684.1; -.
KW   Mitochondrion.
SQ   SEQUENCE   380 AA;  42348 MW;  2500441E CRC32;


  Query Match             5.6%;  Score 108.5;  DB 8;  Length 380;
  Best Local Similarity   20.8%;  Pred. No. 1.2;
  Matches   79;  Conservative   57;  Mismatches  117;  Indels  127;  Gaps   20;

Qy     21 LCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFL 80
          :|::| :||           | |: | :| ||                ||   ||
Db     40 ICLVVQIAT----------GLLMATHYTADT----------------TLAFSSV-------- 68

Qy     81 VVLYKYRCYKFIHGWLI----MSSLMLLFLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFG 136
              : |    :||||      :   : |:  |:::|  |  |   ::  |     || |
Db     68 ----AHTCRNVQYGWLIRNLHANGASMFFICIYLHIGRGL-YYGSYLNKET-----WNTG 117

Qy    137 AVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVILGAISVYDLVAVLCPKG 196
                          :|::  |||  |: |:  |       ||  : :| : :   |
Db    118 ------------------VILLLTLMATAFVGYVLPWGQMSFWGATVITNLFSAIPYIG 158

Qy    197 PLRMLVETAQ---ERNEP----------IFPALIYSSAMVWTVGMAKLDPSSQGAL---- 240
            : :|| |       : |          : | || | ||:    : :   ::  |
Db    159 --QTIVEWAWGGFSVDNPTLTRFFALHFLLPFLIASLAMIHLTFLHETGSNNPLGLTSKC 216

Qy    240 -QLPYDPEMEEDSYDSFG-------------EPSY---PEVFEP--PLTGYPGEELEEE 279
           ::|: |       |: |               |:    || | |  ||   |  : |
Db    217 DKIPFHPYF--SLKDALGFAMMLLFLTTLALFSPNLLGDPENFSPANPLVTPPHIKPE-- 273

Qy    280 EERGVKLGLGDFIF-YSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALP 338
                  |:| |::|       |       ||   ::|:   : :| ::  ::|:
Db    273 ---------WYFLFAYAILRSIPNKLGG-----VLALAASVLVLFLMPILHMSK-QRAMT 317

Qy    339 ALPISITFGLIFYFSTDNLV 358
            |:|    |:|:    ||:
Db    318 FRPLS---QLLFWALVANLL 334


RESULT  12
Q9Z577
ID   Q9Z577      PRELIMINARY;      PRT;   748 AA.
AC   Q9Z577;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   PUTATIVE INTEGRAL MEMBRANE EXPORT PROTEIN.
GN   SC8D9.14.
OS   Streptomyces coelicolor.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   MURPHY L., HARRIS D.;
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
```

```
  Query Match             5.5%;  Score 106.5;  DB 8;  Length 382;
  Best Local Similarity   20.5%;  Pred. No. 1.6;
  Matches   79;  Conservative   53;  Mismatches  115;  Indels  139;  Gaps   18;

Qy      21 LCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFL 80
           :|:|: : |           |  :   :| ||                        :| |
Db      39 MCLIIQILT----------GLFLAMHYTSDT---------------------LTAFS 64

Qy      81 VVLYKYRCYKFIHGWLIM------SSLMLLFLFTYIYLGEVLKTYNVAMDYPTLLLTVWN 134
            |   : |    :||||       :|:  : || ::  |    :|        |    ||
Db      65 SV--AHICRDVNYGWLIRNIHANGASMFFMCLFLHVGRGIYYGSY--------LYKETWN 114

Qy     135 FGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVILGAISVYDLVAVLCP 194
           |                   :|::  :||  |: |:  |      ||  : :|:: :
Db     115 IG------------------VILLLTVMATAFVGYVLPWGQMSFWGATVITNLLSAIPY 155

Qy     195 KGPLRMLVE-------------TAQERNEPIFPALIYSSAMVWTVGMAK--------LDP 233
           |    |||              |       | | :| :  :|  : : :         |||
Db     156 IG--NTLVEWIWGGFSVDKATLTRFFAFHFILPFIILAMVVVHLLFLHETGSNNPTGLDP 213

Qy     234 SSQGALQLPYDP-----------------EMEEDSYDSFGEPSYPEVFEP--PLTGYPG 273
           :|    ::|: |                 :   | |  |:   |: | |  ||   |
Db     214 NSD---KIPFHPYYTIKDILGLFLMIIILLSLAMFSPDLLGD---PDNFTPANPLNTPPH 267

Qy     274 EELEEEEERGVKLGLGDFIF-YSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAV 332
           : |             |:| |::|       |       ||   :||| | : ||  :
Db     268 IKPE----------WYFLFAYAILRSIPNKLGG-----VLALLASILILLIMPLLHTST 311

Qy     333 FKKALPALPISITFGLIFYFSTDNLV 358
           ::::   ||| |   :|:  | ||:
Db     312 -QRSMMFRPISQT---LFWMLTANLI 333


RESULT  15
O35294
ID   O35294      PRELIMINARY;      PRT;   318 AA.
AC   O35294;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   INTEGRIN-ASSOCIATED PROTEIN FORM 4.
GN   IAP.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RA   HUANG A.M., LEE E.H.Y.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF017437; AAB70273.1; -.
KW   Integrin.
SQ   SEQUENCE   318 AA;  34776 MW;  FAA078CB CRC32;


  Query Match             5.5%;  Score 106.5;  DB 11;  Length 318;
  Best Local Similarity   23.0%;  Pred. No. 1.3;
  Matches   40;  Conservative   35;  Mismatches   50;  Indels   49;  Gaps    8;

Qy       3 LTLKYGAKH-----VIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQ 57
           ||||| : |     :::|   : | :||||             | ::: |         |:
Db     158 LTLKYKSSHTNKRIILLLVAGLALTLIVVV------------GAILFIP--------GE 196

Qy      58 RLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLK 117
           : : :       : :||: |  |::| :|  :    |   |:|  :  | | : || ||
Db     197 KPVKNASG----LGLIVISTGILILL-QYNVFMTAFG---MTSFTIAILITQV-LGYVLA 247

Qy     118 TYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYL 171
           : :                | |    |||::    :| :: |: ||::|::
Db     248 VVGMCLCI-------------MACEPVHGPLLISGLGIIALAELLGLVYMKFV 287
```

```
Search completed: March 18, 2000, 22:07:45
Job time: 125 sec
```

```
OM of: US-08-509-359B-137 to: GenEmbl:*   out_format : pfs

Date: Mar 18, 2000  2:37 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-1998 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US08509359/runat_17032000_164055_19467/app_query.fasta.1
 -DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -ALIGN=15 -MODE=LOCAL
 -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08509359
 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-08-509-359B-137
Query length: 448
Database: GenEmbl:*
Database sequences: 821193
Database length: -1518192014
Search time (sec): 533.740000


score_list:
Sequence            Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_pat:AR060156      + 2336.00 3900.72  3.0e-209  2285  ! AR060156 Sequence 136 from pat
gb_pr3:HUME51P       + 2336.00 3900.98  2.9e-209  2229  ! L44577 Homo sapiens S182 prote
gb_pr2:HUMSTM2R      + 2328.00 3887.50  1.6e-208  2236  ! L43964 Homo sapiens (clone F-T
gb_om:AF038937       + 2262.50 3778.58  1.9e-202  2004  ! AF038937 Bos taurus presenilin
gb_ro:MMU57324       + 2232.00 3727.60  1.3e-199  1954  ! U57324 Mus musculus PS-2 mRNA,
gb_ro:AF038935       + 2228.50 3724.54  2.0e-199  1490  ! AF038935 Mus musculus presenil
gb_ro:D83700         + 2227.00 3723.07  2.4e-199  1347  ! D83700 Rattus norvegicus cDNA
gb_ro:RNX99267       + 2224.00 3713.47  8.1e-199  2088  ! X99267 R.norvegicus mRNA for p
gb_pr2:MMPRESEN2     + 2218.00 3708.04  1.6e-198  1335  ! Y10140 M.murinus mRNA for pres
gb_pr3:HSU34349      + 2215.50 3701.29  3.9e-198  1705  ! U34349 Human seven trans-membr
gb_ro:AB004454       + 2213.00 3699.04  5.1e-198  1414  ! AB004454 Rattus norvegicus mRN
gb_ov:D84428         + 1624.50 2704.30  1.3e-142  2490  ! D84428 Xenopus laevis mRNA for
gb_pat:A63557        + 1468.00 2440.57  6.4e-128  2681  ! A63557 Sequence 10 from Patent
gb_pat:AR060155      + 1468.00 2443.82  4.2e-128  1962  ! AR060155 Sequence 135 from pat
gb_ro:MUSS1PR        + 1468.00 2443.81  4.3e-128  1964  ! L42177 Mus musculus S182 prote
gb_pat:AR060154      + 1467.00 2438.47  8.4e-128  2791  ! AR060154 Sequence 133 from pat
gb_pr2:HUMS182R      + 1467.00 2438.57  8.3e-128  2765  ! L42110 Homo sapiens (clone cc3
gb_sts:G27112        + 1467.00 2438.57  8.3e-128  2765  ! G27112 human STS SHGC-31609, s
gb_pat:AR060024      + 1461.00 2428.39  3.1e-127  2791  ! AR060024 Sequence 1 from paten
gb_ro:RATS182        + 1459.50 2433.00  1.7e-127  1407  ! D82578 Rat mRNA for S182 prote
gb_ov:D84427         + 1454.50 2418.17  1.1e-126  2607  ! D84427 Xenopus laevis mRNA for
gb_pr2:HSU40379      + 1454.00 2423.87  5.5e-127  1392  ! U40379 Human presenilin I-463
gb_ro:D82363         + 1453.50 2422.91  6.2e-127  1407  ! D82363 Rattus norvegicus mRNA
gb_pat:AR060026      + 1448.00 2405.49  5.8e-126  3087  ! AR060026 Sequence 5 from paten
gb_pr2:HUMPS1MRNA    + 1448.00 2405.60  5.7e-126  3056  ! L76517 Homo sapiens (clone cc4
gb_pr2:MMPRSNLNI     + 1447.00 2411.82  2.6e-126  1430  ! Z71333 M.murinus presenilin I
gb_om:AF038936       + 1438.50 2394.60  2.3e-125  1896  ! AF038936 Bos taurus presenilin
gb_pr2:HSU40380      + 1433.00 2387.91  5.5e-125  1484  ! U40380 Human presenilin I-374
gb_pat:AR060025      + 1418.00 2359.98  2.0e-123  1929  ! AR060025 Sequence 3 from paten
gb_ov:DRE132931      + 1407.50 2339.94  2.6e-122  2428  ! AJ132931 Danio rerio mRNA for
gb_ov:CCPS1          + 1355.00 2258.71  8.7e-118  1241  ! Y17128 Cyprinus carpio mRNA fo
gb_in1:DMU78084      + 1173.50 1950.08  1.3e-100  1764  ! U78084 Drosophila melanogaster
gb_pat:AR060183      + 1166.50 1937.57  6.7e-100  1895  ! AR060183 Sequence 165 from pat
gb_in1:DMU77934      + 1166.50 1938.06  6.3e-100  1809  ! U77934 Drosophila melanogaster
gb_in2:AF017024      + 1166.50 1937.58  6.7e-100  1894  ! AF017024 Drosophila melanogast
gb_in2:AF171064      + 1043.00 1733.70  1.5e-88   1335  ! AF171064 Caenorhabditis elegan
gb_in1:CEU35660      + 1015.50 1686.56  6.4e-86   1461  ! U35660 Caenorhabditis elegans
gb_in2:AF093402      +  970.50 1596.49  6.7e-81   5858  ! AF093402 Drosophila melanogast
gb_htg5:AC014022     -  967.50 1562.38  5.3e-79  95597  ! AC014022 Drosophila melanogas
gb_in2:AF084184      +  967.00 1599.06  4.8e-81   2601  ! AF084184 Drosophila melanogast
gb_in2:CELF35H12     +  820.50 1328.50  5.6e-66  27102  ! U41540 Caenorhabditis elegans
gb_ro:MMU57325       +  774.00 1283.30  1.9e-63   1146  ! U57325 Mus musculus PS-2short
gb_in2:AF017025      +  576.50  953.46  4.4e-45    944  ! AF017025 Drosophila melanoga
gb_pl1:AP000836      -  567.00  882.27  4.0e-41 190014  ! AP000836 Oryza sativa gen
gb_htg4:AC009840     +  545.00  851.00  2.2e-39 109951  ! AC009840 Drosophila melan

seq_name: gb_pat:AR060156

seq_documentation_block:
LOCUS       AR060156     2285 bp    DNA              PAT       29-SEP-1999
DEFINITION  Sequence 136 from patent US 5840540.
ACCESSION   AR060156
VERSION     AR060156.1  GI:5986606
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2285)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Nucleic acids encoding presenilin II
  JOURNAL   Patent: US 5840540-A 136 24-NOV-1998;
FEATURES             Location/Qualifiers
     source          1. .2285
                     /organism="unknown"
BASE COUNT      537 a    579 c    633 g    521 t     15 others
ORIGIN


alignment_scores:
           Quality: 2336.00            Length:    448
             Ratio:  5.214               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-509-359B-137 x AR060156   ..

 Align seg 1/1  to: AR060156  from: 1  to: 2285

       1 MetLeuThrPheMetAlaSerAspSerGluGluGluValCysAspGluAr 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     366 ATGCTCACATTCATGGCCTCTGACAGCGAGGAAGAAGTGTGTGATGAGCG 415

      17 gThrSerLeuMetSerAlaGluSerProThrProArgSerCysGlnGluG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     416 GACGTCCCTAATGTCGGCCGAGAGCCCCACGCCGCGCTCCTGCCAGGAGG 465

      34 lyArgGlnGlyProGluAspGlyGluAsnThrAlaGlnTrpArgSerGln 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     466 GCAGGCAGGGCCCAGAGGATGGAGAGAATACTGCCCAGTGGAGAAGCCAG 515

      51 GluAsnGluGluAspGlyGluGluAspProAspArgTyrValCysSerGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     516 GAGAACGAGGAGGACGGTGAGGAGGACCCTGACCGCTATGTCTGTAGTGG 565

      67 yValProGlyArgProProGlyLeuGluGluGluLeuThrLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     566 GGTTCCCGGGCGGCCGCCAGGCCTGGAGGAAGAGCTGACCCTCAAATACG 615

      84 lyAlaLysHisValIleMetLeuPheValProValThrLeuCysMetIle 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     616 GAGCGAAGCATGTGATCATGCTGTTTGTGCCTGTCACTCTGTGCATGATC 665

     101 ValValValAlaThrIleLysSerValArgPheTyrThrGluLysAsnGl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     666 GTGGTGGTAGCCACCATCAAGTCTGTGCGCTTCTACACAGAGAAGAATGG 715

     117 yGlnLeuIleTyrThrProPheThrGluAspThrProSerValGlyGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     716 ACAGCTCATCTACACGCCATTCACTGAGGACACACCCTCGGTGGGCCAGC 765

     134 rgLeuLeuAsnSerValLeuAsnThrLeuIleMetIleSerValIleVal 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     766 GCCTCCTCAACTCCGTGCTGAACACCCTCATCATGATCAGCGTCATCGTG 815

     151 ValMetThrIlePheLeuValValLeuTyrLysTyrArgCysTyrLysPh 167
```

```
 17 gThrSerLeuMetSerAlaGluSerProThrProArgSerCysGlnGluG 34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
416 GACGTCCCTAATGTCGGCCGAGAGCCCCACGCCGCGCTCCTGCCAGGAGG 465

 34 lyArgGlnGlyProGluAspGlyGluAsnThrAlaGlnTrpArgSerGln 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
466 GCAGGCAGGGCCCAGAGGATGGAGAGAATACTGCCCAGTGGAGAAGCCAG 515

 51 GluAsnGluGluAspGlyGluGluAspProAspArgTyrValCysSerGl 67
    ||||||||||||||||||||||||||||||||||||||||||||||||||
516 GAGAACGAGGAGGACGGTGAGGAGGACCCTGACCGCTATGTCTGTAGTGG 565

 67 yValProGlyArgProProGlyLeuGluGluGluLeuThrLeuLysTyrG 84
    ||||||||||||||||||||||||||||||||||||||||||||||||||
566 GGTTCCCGGGCGGCCGCCAGGCCTGGAGGAAGAGCTGACCCTCAAATACG 615

 84 lyAlaLysHisValIleMetLeuPheValProValThrLeuCysMetIle 100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
616 GAGCGAAGCATGTGATCATGCTGTTTGTGCCTGTCACTCTGTGCATGATC 665

101 ValValValAlaThrIleLysSerValArgPheTyrThrGluLysAsnGl 117
    ||||||||||||||||||||||||||||||||||||||||||||||||||
666 GTGGTGGTAGCCACCATCAAGTCTGTGCGCTTCTACACAGAGAAGAATGG 715

117 yGlnLeuIleTyrThrProPheThrGluAspThrProSerValGlyGlnA 134
    ||||||||||||||||||||||||||||||||||||||||||||||||||
716 ACAGCTCATCTACACGCCATTCACTGAGGACACACCCTCGGTGGGCCAGC 765

134 rgLeuLeuAsnSerValLeuAsnThrLeuIleMetIleSerValIleVal 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
766 GCCTCCTCAACTCCGTGCTGAACACCCTCATCATGATCAGCGTCATCGTG 815

151 ValMetThrIlePheLeuValValLeuTyrLysTyrArgCysTyrLysPh 167
    ||||||||||||||||||||||||||||||||||||||||||||||||||
816 GTTATGACCATCTTCTTGGTGGTGCTCTACAAGTACCGCTGCTACAAGTT 865

167 eIleHisGlyTrpLeuIleMetSerSerLeuMetLeuLeuPheLeuPheT 184
    ||||||||||||||||||||||||||||||||||||||||||||||||||
866 CATCCATGGCTGGTTGATCATGTCTTCACTGATGCTGCTGTTCCTCTTCA 915

184 hrTyrIleTyrLeuGlyGluValLeuLysThrTyrAsnValAlaMetAsp 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
916 CCTATATCTACCTTGGGGAAGTGCTCAAGACCTACAATGTGGCCATGGAC 965

201 TyrProThrLeuLeuLeuThrValTrpAsnPheGlyAlaValGlyMetVa 217
    ||||||||||||||||||||||||||||||||||||||||||||||||||
966 TACCCCACCCTCTTGCTGACTGTCTGGAACTTCGGGGCAGTGGGCATGGT 1015

217 lCysIleHisTrpLysGlyProLeuValLeuGlnGlnAlaTyrLeuIleM 234
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1016 GTGCATCCACTGGAAGGGCCCTCTGGTGCTGCAGCAGGCCTACCTCATCA 1065

234 etIleSerAlaLeuMetAlaLeuValPheIleLysTyrLeuProGluTrp 250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1066 TGATCAGTGCGCTCATGGCCCTAGTGTTCATCAAGTACCTCCCAGAGTGG 1115

251 SerAlaTrpValIleLeuGlyAlaIleSerValTyrAspLeuValAlaVa 267
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1116 TCCGCGTGGGTCATCCTGGGCGCCATCTCTGTGTATGATCTCGTGGCTGT 1165

257 lLeuCysProLysGlyProLeuArgMetLeuValGluThrAlaGlnGluA 284
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1166 GCTGTGTCCCAAAGGGCCTCTGAGAATGCTGGTAGAAACTGCCCAGGAGA 1215

284 rgAsnGluProIlePheProAlaLeuIleTyrSerSerAlaMetValTrp 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1216 GAAATGAGCCCATATTCCCTGCCCTGATATACTCATCTGCCATGGTGTGG 1265

301 ThrValGlyMetAlaLysLeuAspProSerSerGlnGlyAlaLeuGlnLe 317
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1266 ACGGTTGGCATGGCGAAGCTGGACCCCTCCTCTCAGGGTGCCCTCCAGCT 1315

317 uProTyrAspProGluMetGluGluAspSerTyrAspSerPheGlyGluP 334
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1316 CCCCTACGACCCGGAGATGGAAGAAGACTCCTATGACAGTTTTGGGGAGC 1365

334 roSerTyrProGluValPheGluProProLeuThrGlyTyrProGlyGlu 350
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1366 CTTCATACCCCGAAGTCTTTGAGCCTCCCTTGACTGGCTACCCAGGGGAG 1415

351 GluLeuGluGluGluGluGluArgGlyValLysLeuGlyLeuGlyAspPh 367
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1416 GAGCTGGAGGAAGAGGAGGAAAGGGGCGTGAAGCTTGGCCTCGGGGACTT 1465

367 eIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySerGlyA 384
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1466 CATCTTCTACAGTGTGCTGGTGGGCAAGGCGGCTGCCACGGGCAGCGGGG 1515

384 spTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLeuCys 400
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1516 ACTGGAATACCACGCTGGCCTGCTTCGTGGCCATCCTCATTGGCTTGTGT 1565

401 LeuThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAlaLeuPr 417
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1566 CTGACCCTCCTGCTGCTTGCTGTGTTCAAGAAGGCGCTGCCCGCCCTCCC 1615

417 oIleSerIleThrPheGlyLeuIlePheTyrPheSerThrAspAsnLeuV 434
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1616 CATCTCCATCACGTTCGGGCTCATCTTTTACTTCTCCACGGACAACCTGG 1665

434 alArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIle 448
    |||||||||||||||||||||||||||||||||||||||||||
1666 TGCGGCCGTTCATGGACACCCTGGCCTCCCATCAGCTCTACATC 1709
```

seq_name: gb_pr2:HUMSTM2R

seq_documentation_block:

```
LOCUS       HUMSTM2R     2236 bp    mRNA            PRI       23-AUG-1995
DEFINITION  Homo sapiens (clone F-T03796) STM-2 mRNA, complete cds.
ACCESSION   L43964
VERSION     L43964.1  GI:951202
KEYWORDS    Alzheimer's disease.
SOURCE      Homo sapiens (clone: F-T03796) cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2236)
  AUTHORS   Levy-Lahad,E., Wasco,W., Poorkaj,P., Romano,D.M., Oshima,J.,
            Pettingell,W.H., Yu,C.-E., Jondro,P.D., Schmidt,S.D., Wang,K.,
            Crowley,A.C., Fu,Y.-H., Guenette,S.Y., Galas,D., Nemens,E.,
            Wijsman,E.M., Bird,T.D., Schellenberg,G.D. and Tanzi,R.E.
  TITLE     Candidate gene for the chromosome 1 familial Alzheimer's disease
            locus
  JOURNAL   Science 269 (5226), 973-977 (1995)
  MEDLINE   95365816
FEATURES             Location/Qualifiers
     source          1. .2236
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="F-T03796"
     gene            368. .1714
                     /gene="STM-2"
     CDS             368. .1714
                     /gene="STM-2"
                     /codon_start=1
                     /protein_id="AAB59557.1"
                     /db_xref="GI:951203"
                     /translation="MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENT
                     AQWRSQENEEDGEEDPDRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVV
                     VATIKSVRFYTEKNGQLIYTTFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLY
                     KYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVC
                     IHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRM
                     LVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQLPYDPEMEEDSYDSFGEP
                     SYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACF
                     VAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI"
     allele          789
```

```
  JOURNAL   Submitted (16-DEC-1997) Molecular Biology, Tokyo Institute of
            Psychiatry, 2-1-8 Kami-Kitazawa, Setagaya-ku, Tokyo 156, Japan
FEATURES             Location/Qualifiers
     source          1. .2004
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /tissue_type="brain"
     CDS             83. .1432
                     /note="causal gene for familial Alzheimer's disease"
                     /codon_start=1
                     /product="presenilin 2"
                     /protein_id="AAD39024.1"
                     /db_xref="GI:5070118"
                     /translation="MLTFMASDSEEEVCDERTSLMSAESPTPRSCQDGRQGLEDGESA
                     AQWRSQESEEDHEEEDPDRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIV
                     VVATIKSVRFYTEKNGQLIYTPFSEDTPSVGQRLLNSVLNTLIMISVIVTMTIFLVVL
                     YKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYNVAMDYPTLFLTVWNFGAVGMV
                     CIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLR
                     MLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQLPYDPEMEEDSYDSFGE
                     PSYPDVFEPPLPGYPGEELEEEEERGVKLGLGDFIFYSVLVGKAAAMGSGDWNTTLAC
                     FVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI"
     polyA_signal    1982. .1987
BASE COUNT      395 a    573 c    596 g    440 t
ORIGIN


alignment_scores:
            Quality: 2262.50              Length:    449
              Ratio:  5.096                Gaps:      1
 Percent Similarity:  98.886   Percent Identity: 97.105

alignment_block:
 US-08-509-359B-137 x AF038937   ..

 Align seg 1/1  to: AF038937  from: 1  to: 2004

       1 MetLeuThrPheMetAlaSerAspSerGluGluGluValCysAspGluAr 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      83 ATGCTCACATTCATGGCCTCTGACAGCGAGGAAGAAGTGTGTGATGAGCG 132

      17 gThrSerLeuMetSerAlaGluSerProThrProArgSerCysGlnGluG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||:::|
     133 GACGTCCCTGATGTCAGCCGAGAGCCCTACGCCGCGCTCCTGCCAGGACG 182

      34 lyArgGlnGlyProGluAspGlyGluAsnThrAlaGlnTrpArgSerGln 50
         ||||||||||||   ||||||||||||:::::::||||||||||||||||
     183 GCCGGCAGGGCCTGGAGGACGGAGAGAGTGCCGCCCAGTGGAGAAGCCAG 232

      51 GluAsnGluGluAsp...GlyGluGluAspProAspArgTyrValCysSe 66
         |||:::|||||||||      ||||||||||||||||||||||||||||||
     233 GAGAGTGAGGAGGACCACGAGGAGGAGGACCCTGACCGCTACGTCTGCAG 282

      66 rGlyValProGlyArgProProGlyLeuGluGluGluLeuThrLeuLysT 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     283 TGGGGTTCCTGGGCGGCCGCCAGGCCTGGAGGAGGAGCTGACCCTCAAAT 332

      83 yrGlyAlaLysHisValIleMetLeuPheValProValThrLeuCysMet 99
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     333 ACGGGGCAAAGCATGTGATCATGTTGTTTGTGCCTGTCACACTGTGCATG 382

     100 IleValValValAlaThrIleLysSerValArgPheTyrThrGluLysAs 116
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     383 ATCGTGGTGGTGGCCACCATCAAGTCCGTGCGCTTCTACACAGAGAAGAA 432

     116 nGlyGlnLeuIleTyrThrProPheThrGluAspThrProSerValGlyG 133
         ||||||||||||||||||||||:::|||||||||||||||||||||||||
     433 TGGACAGCTCATCTATACCCCGTTCTCCGAGGACACGCCCTCAGTGGGCC 482

     133 lnArgLeuLeuAsnSerValLeuAsnThrLeuIleMetIleSerValIle 149
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     483 AGCGGCTCCTCAACTCCGTGCTCAACACCCTCATCATGATCAGCGTCATT 532

     150 ValValMetThrIlePheLeuValValLeuTyrLysTyrArgCysTyrLy 166
         |||:::||||||||||||||||||||||||||||||||||||||||||||
     533 GTCACCATGACGATCTTCCTGGTTGTGCTCTACAAGTACCGCTGCTACAA 582

     166 sPheIleHisGlyTrpLeuIleMetSerSerLeuMetLeuLeuPheLeuP 183
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     583 GTTCATCCACGGTTGGCTGATCATGTCCTCTCTGATGTTGCTCTTCCTCT 632

     183 heThrTyrIleTyrLeuGlyGluValLeuLysThrTyrAsnValAlaMet 199
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     633 TCACCTACATCTACCTCGGGGAAGTGCTCAAGACCTACAATGTGGCCATG 682

     200 AspTyrProThrLeuLeuLeuThrValTrpAsnPheGlyAlaValGlyMe 216
         ||||||||||||||:::||||||||||||||||||||||||||||||||
     683 GACTACCCCACCCTGTTCCTGACCGTCTGGAACTTCGGGGCGGTGGGCAT 732

     216 tValCysIleHisTrpLysGlyProLeuValLeuGlnGlnAlaTyrLeuI 233
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     733 GGTGTGCATCCACTGGAAGGGCCCCCTAGTGCTGCAGCAGGCCTACCTCA 782

     233 leMetIleSerAlaLeuMetAlaLeuValPheIleLysTyrLeuProGlu 249
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     783 TCATGATCAGCGCGCTCATGGCCTTGGTGTTCATCAAGTACCTCCCGGAG 832

     250 TrpSerAlaTrpValIleLeuGlyAlaIleSerValTyrAspLeuValAl 266
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     833 TGGTCCGCCTGGGTCATTCTGGGCGCCATCTCCGTGTATGATCTTGTGGC 882

     266 aValLeuCysProLysGlyProLeuArgMetLeuValGluThrAlaGlnG 283
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     883 TGTGCTGTGTCCCAAAGGGCCACTGAGGATGCTCGTGGAAACCGCGCAGG 932

     283 luArgAsnGluProIlePheProAlaLeuIleTyrSerSerAlaMetVal 299
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     933 AGAGAAATGAGCCCATATTCCCTGCCCTCATATACTCATCCGCCATGGTG 982

     300 TrpThrValGlyMetAlaLysLeuAspProSerSerGlnGlyAlaLeuGl 316
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     983 TGGACCGTGGGCATGGCCAAGCTGGACCCTTCCTCTCAGGGAGCCCTTCA 1032

     316 nLeuProTyrAspProGluMetGluGluAspSerTyrAspSerPheGlyG 333
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1033 GCTCCCCTACGACCCGGAGATGGAAGAAGACTCCTATGACAGTTTTGGGG 1082

     333 luProSerTyrProGluValPheGluProProLeuThrGlyTyrProGly 349
         ||||||||||||||:::|||||||||||||||||   ||||||||||||
    1083 AGCCCTCGTACCCTGACGTCTTTGAGCCCCCGCTGCCTGGCTACCCGGGG 1132

     350 GluGluLeuGluGluGluGluGluArgGlyValLysLeuGlyLeuGlyAs 366
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1133 GAGGAGCTGGAGGAAGAGGAGGAAAGGGGCGTGAAGCTCGGCCTCGGAGA 1182

     366 pPheIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySerG 383
         ||||||||||||||||||||||||||||||||||||||   |||||||||
    1183 CTTCATCTTCTACAGTGTGCTGGTGGGCAAGGCGGCCGCCATGGGCAGCG 1232

     383 lyAspTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLeu 399
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1233 GGGACTGGAACACCACACTGGCCTGCTTTGTGGCCATCCTCATTGGTTTG 1282

     400 CysLeuThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAlaLe 416
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1283 TGTCTGACCCTCCTGCTGCTCGCCGTGTTCAAGAAGGCGCTGCCCGCCCT 1332

     416 uProIleSerIleThrPheGlyLeuIlePheTyrPheSerThrAspAsnL 433
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1333 CCCCATCTCCATCACGTTCGGGCTCATCTTCTACTTCTCCACGGACAACC 1382

     433 euValArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIle 448
         |||||||||||||||||||||||||||||||||||||||||||||||||
    1383 TGGTGCGGCCTTTCATGGACACCCTGGCCTCCCATCAGCTCTACATC 1429

seq_name: gb_ro:MMU57324
```

```
      |||||||||||||:::|||||||   ||||||   ||||||||||||||||
   1042 CCTCATACCCTGAAGCCTTCGAAGCCCCCCTGCCTGGCTACCCAGGGGAG 1091

    351 GluLeuGluGluGluGluGluArgGlyValLysLeuGlyLeuGlyAspPh 367
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
   1092 GAGCTGGAGGAGGAGGAGGAAAGGGGCGTGAAGCTCGGCCTGGGAGACTT 1141

    367 eIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySerGlyA 384
      |||||||||||||||||||||||||||||||||||||||||||:::||||
   1142 CATCTTCTACAGCGTCCTGGTGGGCAAGGCTGCAGCCACTGGCAACGGAG 1191

    384 spTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLeuCys 400
      |||||||||||||||||||||||||:::|||||||||||||||||||||||
   1192 ACTGGAACACTACGCTGGCCTGTTTTATCGCCATCCTCATTGGCTTGTGT 1241

    401 LeuThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAlaLeuPr 417
      |||||||||||||||||||||||||||||||||||||||||||||||||||
   1242 CTCACCCTCCTGCTGCTTGCTGTGTTCAAGAAGGCTCTGCCCGCCCTCCC 1291

    417 oIleSerIleThrPheGlyLeuIlePheTyrPheSerThrAspAsnLeuV 434
      |||||||||||||||||||||||||||||||||||||||||||||||||||
   1292 CATCTCCATCACCTTTGGACTCATCTTCTACTTCTCCACAGACAACCTGG 1341

    434 alArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIle 448
      |||||||||||||||||||||||||||||||||||||||||||
   1342 TGCGCCCTTTCATGGACACTCTGGCCTCCCACCAGCTCTACATC 1385

seq_name: gb_ro:AF038935

seq_documentation_block:
LOCUS       AF038935     1490 bp    mRNA            ROD       28-DEC-1997
DEFINITION  Mus musculus presenilin 2 mRNA, complete cds.
ACCESSION   AF038935
VERSION     AF038935.1  GI:2724097
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1490)
  AUTHORS   Sahara,N., Mori,H. and Shirasawa,T.
  TITLE     Molecular cloning of mouse presenilin 2 gene
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1490)
  AUTHORS   Sahara,N., Mori,H. and Shirasawa,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-1997) Molecular Biology, Tokyo Institute of
            Psychiatry, 2-1-8 Kami-Kitazawa, Setagaya-ku, Tokyo 156, Japan
FEATURES             Location/Qualifiers
     source          1. .1490
                     /organism="Mus musculus"
                     /strain="NIH SwissMus"
                     /db_xref="taxon:10090"
     CDS             19. .1362
                     /note="causal gene for familial Alzheimer's disease"
                     /codon_start=1
                     /product="presenilin 2"
                     /protein_id="AAB92660.1"
                     /db_xref="GI:2724098"
                     /translation="MLAFMASDSEEEVCDERTSLMSAESPTSRSCQEGRPGPEDGEST
                     AQWRTQESEEDCEEDPDRYACSGAPGRPSGLEEELTLKYGAKHVIMLFVPVTLCMIVV
                     VATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLY
                     KYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYNVAMDYPTLFLAVWNFGAVGMVC
                     IHWKGPLVLQQAYLIVISALMALVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRM
                     LVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQLPYDPEMEDSYDSFGEPS
                     YPEAFEAPLPGYPGEELEEEEERGVKLGLGDFIFYSVLVGKAAATGNGDWNTTLACFI
                     AILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI"
BASE COUNT      315 a    418 c    416 g    341 t
ORIGIN


alignment_scores:
          Quality: 2228.50            Length:    448
```

```
            Ratio:  5.053              Gaps:      1
 Percent Similarity:  98.438   Percent Identity: 95.759

alignment_block:
 US-08-509-359B-137 x AF038935   ..

 Align seg 1/1  to: AF038935  from: 1  to: 1490

       1 MetLeuThrPheMetAlaSerAspSerGluGluGluValCysAspGluAr 17
         ||||||:::||||||||||||||||||||||||||||||||||||||||||
      19 ATGCTCGCATTCATGGCCTCTGACAGCGAGGAAGAGGTGTGTGATGAGCG 68

      17 gThrSerLeuMetSerAlaGluSerProThrProArgSerCysGlnGluG 34
         |||||||||||||||||||||||||||||||   ||||||||||||||||
      69 GACGTCCTTGATGTCAGCCGAGAGCCCCACATCTCGCTCCTGCCAGGAAG 118

      34 lyArgGlnGlyProGluAspGlyGluAsnThrAlaGlnTrpArgSerGln 50
         |||||   ||||||||||||||||||:::|||||||||||||||:::|||
     119 GCAGGCCAGGCCCGGAGGATGGAGAGAGCACTGCCCAGTGGAGGACCCAG 168

      51 GluAsnGluGluAspGlyGluGluAspProAspArgTyrValCysSerGl 67
         |||:::||||||||||   ||||||||||||||||||||||:::||||||||
     169 GAGAGCGAAGAAGACTGTGAAGAGGACCCGGACCGCTACGCATGCAGTGG 218

      67 yValProGlyArgProProGlyLeuGluGluGluLeuThrLeuLysTyrG 84
         |:::|||||||||||   ||||||||||||||||||||||||||||||||
     219 GGCTCCTGGGCGACCGTCGGGCCTGGAGGAAGAGCTGACCCTGAAGTATG 268

      84 lyAlaLysHisValIleMetLeuPheValProValThrLeuCysMetIle 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     269 GGGCGAAGCATGTGATCATGCTATTCGTGCCTGTCACGCTGTGTATGATC 318

     101 ValValValAlaThrIleLysSerValArgPheTyrThrGluLysAsnGl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     319 GTGGTGGTGGCCACTATCAAGTCTGTGCGTTTCTACACTGAGAAGAACGG 368

     117 yGlnLeuIleTyrThrProPheThrGluAspThrProSerValGlyGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     369 GCAGCTCATCTACACGCCCTTCACGGAGGACACGCCCTCGGTGGGCCAGC 418

     134 rgLeuLeuAsnSerValLeuAsnThrLeuIleMetIleSerValIleVal 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     419 GGCTCCTCAACTCTGTGCTGAACACCCTCATCATGATCAGCGTCATTGTA 468

     151 ValMetThrIlePheLeuValValLeuTyrLysTyrArgCysTyrLysPh 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     469 GTCATGACCATCTTTCTCGTGGTACTCTACAAGTATCGATGCTACAAGTT 518

     167 eIleHisGlyTrpLeuIleMetSerSerLeuMetLeuLeuPheLeuPheT 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     519 CATCCATGGCTGGCTGATCATGTCCTCCCTGATGCTCCTCTTCTTGTTCA 568

     184 hrTyrIleTyrLeuGlyGluValLeuLysThrTyrAsnValAlaMetAsp 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     569 CCTACATCTACCTCGGGGAAGTGCTCAAGACCTACAATGTGGCCATGGAC 618

     201 TyrProThrLeuLeuLeuThrValTrpAsnPheGlyAlaValGlyMetVa 217
         ||||||||||||:::|||:::||||||||||||||||||||||||||||||
     619 TATCCCACACTCTTCCTGGCTGTCTGGAACTTCGGGGCAGTGGGCATGGT 668

     217 lCysIleHisTrpLysGlyProLeuValLeuGlnGlnAlaTyrLeuIleM 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||:
     669 GTGCATCCACTGGAAGGGGCCTCTGGTGCTGCAGCAGGCTTATCTTATTG 718

     234 etIleSerAlaLeuMetAlaLeuValPheIleLysTyrLeuProGluTrp 250
         ::||||||||||||||||||||||||||||||||||||||||||||||||||
     719 TGATCAGCGCACTCATGGCCCTGGTGTTCATCAAGTACCTGCCGGAGTGG 768

     251 SerAlaTrpValIleLeuGlyAlaIleSerValTyrAspLeuValAlaVa 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     769 TCTGCCTGGGTCATCTTGGGTGCCATCTCTGTGTACGATCTCGTGGCCGT 818

     267 lLeuCysProLysGlyProLeuArgMetLeuValGluThrAlaGlnGluA 284
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   501 CATCCATGGCTGGCTGATCATGTCCTCCCTGATGCTCCTCTTCTTGTTCA 550

   184 hrTyrIleTyrLeuGlyGluValLeuLysThrTyrAsnValAlaMetAsp 200
       ||||||||||||||||||||||:::|||||||||||||||||||||||||
   551 CCTACATCTACCTCGGGGAAGTGTTCAAGACCTACAATGTGGCCATGGAC 600

   201 TyrProThrLeuLeuLeuThrValTrpAsnPheGlyAlaValGlyMetVa 217
       ||||||||||||:::|||:::|||||||||||||||||||||||||||||
   601 TACCCCACACTGTTCCTGGCTGTCTGGAACTTCGGGGCAGTGGGCATGGT 650

   217 lCysIleHisTrpLysGlyProLeuValLeuGlnGlnAlaTyrLeuIleM 234
       ||||||||||||||||||||||||||||||||||||||||||||||||||:
   651 GTGCATCCACTGGAAGGGTCCCCTGGTGCTCCAGCAGGCTTACCTCATTG 700

   234 etIleSerAlaLeuMetAlaLeuValPheIleLysTyrLeuProGluTrp 250
       ::||||||||||||||||||||||||||||||||||||||||||||||||
   701 TGATCAGCGCACTCATGGCCCTGGTGTTCATCAAGTACCTGCCGGAGTGG 750

   251 SerAlaTrpValIleLeuGlyAlaIleSerValTyrAspLeuValAlaVa 267
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   751 TCCGCTTGGGTCATCCTGGGCGCCATCTCTGTGTACGATCTCGTGGCTGT 800

   267 lLeuCysProLysGlyProLeuArgMetLeuValGluThrAlaGlnGluA 284
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   801 GCTGTGCCCCAAAGGGCCACTGAGGATGCTGGTGGAAACTGCCCAGGAAA 850

   284 rgAsnGluProIlePheProAlaLeuIleTyrSerSerAlaMetValTrp 300
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   851 GAAACGAGCCCATATTTCCTGCCCTCATATACTCATCTGCCATGGTGTGG 900

   301 ThrValGlyMetAlaLysLeuAspProSerSerGlnGlyAlaLeuGlnLe 317
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   901 ACGGTAGGCATGGCGAAGCTGGACCCCTCCTCTCAGGGAGCACTGCAGCT 950

   317 uProTyrAspProGluMetGluGluAspSerTyrAspSerPheGlyGluP 334
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   951 CCCTTACGATCCAGAGATGGAAGAAGACTCCTATGACAGTTTTGGGGAAC 1000

   334 roSerTyrProGluValPheGluProProLeuThrGlyTyrProGlyGlu 350
       ||||||||||||||:::||||||   |||      ||||||||||||||||
  1001 CCTCATACCCTGAAGCCTTCGAAGCTCCCCAGCCTGGCTACCCAGGGGAA 1050

   351 GluLeuGluGluGluGluGluArgGlyValLysLeuGlyLeuGlyAspPh 367
       |||   ||||||||||||||||||||||||||||||||||||||||||||
  1051 GAGCCGGAGGAGGAGGAGGAAAGGGGTGTAAAGCTTGGTCTGGGAGACTT 1100

   367 eIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySerGlyA 384
       ||||||||||||||||||||||||||||||||||||||||||:::||||||
  1101 CATCTTCTACAGCGTTCTGGTGGGCAAGGCAGCAGCCACAGGCAACGGAG 1150

   384 spTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLeuCys 400
       |||||:::|||||||||||||||||:::||||||||||||||||||||||
  1151 ACTGGAGCACGACGCTGGCCTGCTTCATCGCCATCCTCATTGGCTTGTGT 1200

   401 LeuThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAlaLeuPr 417
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1201 CTCACCCTCCTGCTGCTCGCCGTGTTCAAGAAGGCCCTGCCCGCCCTCCC 1250

   417 oIleSerIleThrPheGlyLeuIlePheTyrPheSerThrAspAsnLeuV 434
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1251 CATCTCCATCACCTTCGGGCTCATCTTTTACTTCTCCACAGACAACCTGG 1300

   434 alArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIle 448
       |||||||||||||||||||||||||||||||||||||||||||
  1301 TGCGTCCTTTCATGGACACGTTGGCCTCCCACCAGCTCTACATC 1344

seq_name: gb_ro:RNX99267

seq_documentation_block:
LOCUS       RNX99267     2088 bp    mRNA            ROD       01-AUG-1998
DEFINITION  R.norvegicus mRNA for presenilin-2.
ACCESSION   X99267
VERSION     X99267.1  GI:3378674
KEYWORDS    presenilin-2; pS2 gene.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 2088)
  AUTHORS   Frentzel,S., Abdel,A.S. and Luebbert,H.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2088)
  AUTHORS   Frentzel,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUL-1996) S. Frentzel, Sandoz Pharma Ltd, Preclinical
            Research, PO Box, CH-4002 Basel, Switzerland
FEATURES             Location/Qualifiers
     source          1. .2088
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /tissue_type="total brain"
                     /dev_stage="adult"
     gene            121. .1467
                     /gene="PS2"
     CDS             121. .1467
                     /gene="PS2"
                     /codon_start=1
                     /product="presenilin-2"
                     /protein_id="CAA67663.1"
                     /db_xref="GI:3378675"
                     /db_xref="SPTREMBL:O88777"
                     /translation="MLTFMATDSEEEVCDERTSLMSAESPTSRSCQDSRPGPEDGENT
                     AQWRSQENEDDCEEDPDHYACSGVPGRPSGLEEELTLKYGAKHVIMLFVPVTLCMIVV
                     VATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLY
                     KYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVFKTYNVAMDYPTLFLAVWNFGAVGMVC
                     IHWKGPLVLQQAYLIVISALMALVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRM
                     LVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQLPYDPEMEEDSYDSFGEP
                     SYPEAFEAPQPGYPGEEPEEEEERGVKLGLGDFIFYSVLVGKAAATGNGDWSTTLACF
                     IAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI"
BASE COUNT      483 a    557 c    570 g    478 t
ORIGIN


alignment_scores:
            Quality: 2224.00             Length:    448
              Ratio:  5.055                Gaps:      0
 Percent Similarity:  98.214   Percent Identity: 95.089

alignment_block:
 US-08-509-359B-137 x RNX99267   ..

 Align seg 1/1  to: RNX99267  from: 1  to: 2088

       1 MetLeuThrPheMetAlaSerAspSerGluGluGluValCysAspGluAr 17
         ||||||||||||||||||:::|||||||||||||||||||||||||||||
     121 ATGCTCACATTCATGGCCACTGATAGCGAGGAAGAGGTGTGCGATGAGCG 170

      17 gThrSerLeuMetSerAlaGluSerProThrProArgSerCysGlnGluG 34
         ||||||||||||||||||||||||||||||   ||||||||||||::::
     171 GACGTCCTTGATGTCAGCCGAGAGCCCTACATCACGTTCCTGCCAGGACA 220

      34 lyArgGlnGlyProGluAspGlyGluAsnThrAlaGlnTrpArgSerGln 50
         ::|||   ||||||||||||||||||||||||||||||||||||||||||
     221 GCAGGCCAGGCCCAGAGGATGGAGAGAACACTGCCCAGTGGAGGAGCCAG 270

      51 GluAsnGluGluAspGlyGluGluAspProAspArgTyrValCysSerGl 67
         |||||||||:::|||   ||||||||||||||:::|||:::||||||||
     271 GAGAACGAGGACGACTGTGAGGAGGACCCGGACCACTATGCCTGCAGTGG 320

      67 yValProGlyArgProProGlyLeuGluGluGluLeuThrLeuLysTyrG 84
         ||||||||||||||   ||||||||||||||||||||||||||||||||||
     321 GGTTCCTGGGCGACCATCGGGCCTGGAGGAGGAGCTGACCCTCAAGTATG 370

      84 lyAlaLysHisValIleMetLeuPheValProValThrLeuCysMetIle 100
```

```
  1 MetLeuThrPheMetAlaSerAspSerGluGluGluValCysAspGluAr 17
    ||||||||||||||||||||||||||||||||||||||||||||||||||
  1 ATGCTCACATTCATGGCCTCTGACAGCGAGGAGGAAGTGTGCGACGAGCG 50

 17 gThrSerLeuMetSerAlaGluSerProThrProArgSerCysGlnGluG 34
    ||||||||||||||||||||||||||||:::||||||||||||||||||||
 51 GACGTCCCTGATGTCGGCCGAGAGCCCCTCGCCACGCTCCTGCCAGGAGG 100

 34 lyArgGlnGlyProGluAspGlyGluAsnThrAlaGlnTrpArgSerGln 50
    ||   ||||||||||||||||||:::::|||||||||||||||   |||
101 GCGGGCAGGGCCCGGAGGATGGAGACAGCACTGCCCAGTGGAGAATCCAG 150

 51 GluAsnGluGluAspGlyGluGluAspProAspArgTyrValCysSerGl 67
    ::::::||||||||||||||||||||||||||||||||||||  |||||
151 GATAGTGAGGAGGACGGTGAGGAGGACCCTGACCGCTATGTCAGCAGCGG 200

 67 yValProGlyArgProProGlyLeuGluGluGluLeuThrLeuLysTyrG 84
    ||||||||||||||||||||   ||||||||||||||||||||||||||
201 GGTCCCCGGGCGGCCACCGGGCCCGGAGGAAGAGCTGACCCTCAAATATG 250

 84 lyAlaLysHisValIleMetLeuPheValProValThrLeuCysMetIle 100
    ||||||||||||||||||||   ||||||||||||||||||||||||||
251 GAGCAAAGCATGTGATCATGCTGTCCGTGCCTGTCACCCTGTGCATGATT 300

101 ValValValAlaThrIleLysSerValArgPheTyrThrGluLysAsnGl 117
    ||||||||||||||||||||||||||||||||||||||||||||||||||
301 GTGGTGGTGGCCACCATCAAGTCCGTGCGCTTCTACACGGAGAAGAACGG 350

117 yGlnLeuIleTyrThrProPheThrGluAspThrProSerValGlyGlnA 134
    ||||||||||||||||||||||||||||||||||||||||||:::||||
351 ACAGCTCATCTACACGCCATTCACTGAGGACACACCCTCGGTGAGCCAGC 400

134 rgLeuLeuAsnSerValLeuAsnThrLeuIleMetIleSerValIleVal 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
401 GCCTCCTCAACTCCGTGCTGAACACGCTCATCATGATCAGCGTCATCGTG 450

151 ValMetThrIlePheLeuValValLeuTyrLysTyrArgCysTyrLysPh 167
    ||||||||||||||||||||||||||||||||||||||||||||||||||
451 GTCATGACCATCTTCCTGGTGGTGCTGTACAAGTACCGCTGCTACAAGTT 500

167 eIleHisGlyTrpLeuIleMetSerSerLeuMetLeuLeuPheLeuPheT 184
    ||||||||||||||||||||||||||||||||||||||||||||||||||
501 CATCCACGGCTGGCTGATCATGTCCTCCCTGATGTTGCTGTTCCTCTTCA 550

184 hrTyrIleTyrLeuGlyGluValLeuLysThrTyrAsnValAlaMetAsp 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
551 CCTACATCTACCTCGGGGAAGTGCTCAAGACCTACAACGTGGCCATGGAC 600

201 TyrProThrLeuLeuLeuThrValTrpAsnPheGlyAlaValGlyMetVa 217
    ||||||||||||:::|||||||||||||||||||||||||||||||||||
601 TACCCCACGCTTGTGCTGACCGTCTGGAACTTTGGGGCGGTGGGCATGGT 650

217 lCysIleHisTrpLysGlyProLeuValLeuGlnGlnAlaTyrLeuIleM 234
    ||||||||||||||||||||||||:::||||||||||||||||||||||
651 ATGCATCCACTGGAAGGGGCCTCTCATGCTGCAGCAGGCCTACCTCATCG 700

234 etIleSerAlaLeuMetAlaLeuValPheIleLysTyrLeuProGluTrp 250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
701 CGATCAGCGCGCTCATGGCCCTGGTGTTCATCAAGTACCTCCCGGAGTGG 750

251 SerAlaTrpValIleLeuGlyAlaIleSerValTyrAspLeuValAlaVa 267
    ||||||||||||||||||||||||||||||||||||||||||||||||||
751 TCCGCCTGGGTCATCTTGGGCGCCATCTCTGTGTACGATCTTGTCGCTGT 800

267 lLeuCysProLysGlyProLeuArgMetLeuValGluThrAlaGlnGluA 284
    ||||||||||||||||||||||||||||||||||||||||||||||||||
801 ACTGTGTCCCAAAGGGCCGCTGAGAATGCTGGTGGAAACTGCCCAGGAGA 850

284 rgAsnGluProIlePheProAlaLeuIleTyrSerSerAlaMetValTrp 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
851 GAAACGAGCCCATATTCCCTGCCCTGATATACTCATCTGCCATGGTGTGG 900

301 ThrValGlyMetAlaLysLeuAspProSerSerGlnGlyAlaLeuGlnLe 317
    ||||||||||||||||||||||||||||||||||||||||||||||||||
901 ACAGTGGGCATGGCCAAGCTGGACCCCTCCTCTCAGGGAGCCCTGCAGCT 950

317 uProTyrAspProGluMetGluGluAspSerTyrAspSerPheGlyGluP 334
    ||||||||||||||||||||||||||||||||||||||:::|||||||
951 CCCCTATGACCCTGAGATGGAAGAAGACTCCTATGACAGCTTGGGGGAGC 1000

334 roSerTyrProGluValPheGluProProLeuThrGlyTyrProGlyGlu 350
    ||||||||||||||||||||||   ||||||   ||||||||||||||||
1001 CCTCATACCCCGAAGTCTTTGAGGCCCCCCTGCCTGGCTACCCAGGGGAG 1050

351 GluLeuGluGluGluGluGluArgGlyValLysLeuGlyLeuGlyAspPh 367
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1051 GAGCTGGAGGAAGAGGAGGAAAGGGGCGTGAAGCTCGGCCTCGGGGACTT 1100

367 eIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySerGlyA 384
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1101 CATCTTCTACAGCGTGCTGGTGGGCAAGGCGGCGGCGACGGGCAGCGGGG 1150

384 spTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLeuCys 400
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1151 ACTGGAACACCACCCTGGCCTGCTTCGTGGCCATCCTCATCGGCTTGTGT 1200

401 LeuThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAlaLeuPr 417
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1201 CTGACCCTCCTGCTGCTTGCTGTGTTCAAGAAAGCGCTGCCCGCCCTCCC 1250

417 oIleSerIleThrPheGlyLeuIlePheTyrPheSerThrAspAsnLeuV 434
    ||||||||||||||||||||||:::||||||||||||||||||||||||
1251 CATCTCCATCACGTTCGGGCTCGTCTTCTACTTCTCCACGGATAACCTGG 1300

434 alArgProPheMetAspThrLeuAlaSerHisGln 445
    ||||||||||||||||||||||||   ||||||
1301 TGCGCCCCTTCATGGACACCTTGGCCTACCATCAG 1335
```

seq_name: gb_pr3:HSU34349

seq_documentation_block:

LOCUS HSU34349 1705 bp mRNA PRI 29-NOV-1995
DEFINITION Human seven trans-membrane domain protein (AD3LP/AD5) mRNA, complete cds.
ACCESSION U34349
VERSION U34349.1 GI:1079575
KEYWORDS Alzheimer's disease.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1673)
AUTHORS Li,J., Ma,J. and Potter,H.
TITLE Identification and expression analysis of a potential familial Alzheimer disease gene on chromosome 1 related to AD3
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 92 (26), 12180-12184 (1995)
MEDLINE 96109229
REFERENCE 2 (bases 1 to 1673)
AUTHORS Li,J.
TITLE Direct Submission
JOURNAL Submitted (21-AUG-1995) Jinhe Li, Neurobiology, Harvard Medical School, 220 Longwood Ave., B2-502, Boston, MA 02115, USA
FEATURES Location/Qualifiers
source 1. .1705
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="Contig WC-516, 289 cM from centromere"
/chromosome="1"
5'UTR 1. .91
gene 92. .1420
/gene="AD3LP/AD5"
CDS 92. .1420
/gene="AD3LP/AD5"
/note="A potential familial Alzheimer's disease gene related to AD3; Similar to the Mus musculus product

```
  AUTHORS   Tanahashi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAY-1997) to the DDBJ/EMBL/GenBank databases. Hiroshi
            Tanahashi, National Institute of Neuroscience, Division of
            Demyelinating Disease and Aging; 4-1-1 Ogawahigashi, Kodaira, Tokyo
            187, Japan (E-mail:tanahash@ncnaxp.ncnp.go.jp, Tel:81-423-41-1717,
            Fax:81-423-46-1747)
REFERENCE   2  (sites)
  AUTHORS   Tanahashi,H. and Tabira,T.
  TITLE     Cloning of the cDNA encoding rat presenilin-2
  JOURNAL   Biochim. Biophys. Acta 1396 (3), 259-262 (1998)
  MEDLINE   98207716
COMMENT     Sequence updated (06-Jun-1997)
            Sequence updated (08-Jun-1997).
FEATURES             Location/Qualifiers
     source          1. .1414
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /cell_type="brain"
     CDS             8. .1354
                     /codon_start=1
                     /product="presenilin-2"
                     /protein_id="BAA20406.1"
                     /db_xref="GI:2190043"
                     /translation="MLTFMASDSEEEVCDERTSLMSAESPTSRSCQDSRPGPEDGENT
                     AQWRSQENEDDCEEDPDHYACSGVPGRPSGLEEELTLKYGANDVIMLFVPVTLCMIVV
                     VATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLY
                     KYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVFKTYNVAMDYPTLFLAVWNFGAVGMVC
                     IHWKGPLVLQQAYLIVISALMALVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRM
                     LVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQLPYDPEMEEDSYDSFGEP
                     SYPEAFEAPQPGYPGEEPEEEEERGVKLGLGDFIFYSVLVGKAAATGNGDWSTTLACF
                     IAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI"
BASE COUNT      288 a    410 c    398 g    318 t
ORIGIN


alignment_scores:
            Quality: 2213.00              Length:    448
              Ratio:  5.041                Gaps:      0
 Percent Similarity:  97.991   Percent Identity: 94.866

alignment_block:
 US-08-509-359B-137 x AB004454   ..

 Align seg 1/1  to: AB004454  from: 1  to: 1414

       1 MetLeuThrPheMetAlaSerAspSerGluGluGluValCysAspGluAr 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
       8 ATGCTCACATTCATGGCCTCTGATAGCGAGGAAGAGGTGTGCGATGAGCG 57

      17 gThrSerLeuMetSerAlaGluSerProThrProArgSerCysGlnGluG 34
         |||||||||||||||||||||||||||||||   ||||||||||||::::
      58 GACGTCCTTGATGTCAGCCGAGAGCCCTACATCACGTTCCTGCCAGGACA 107

      34 lyArgGlnGlyProGluAspGlyGluAsnThrAlaGlnTrpArgSerGln 50
         ::|||   ||||||||||||||||||||||||||||||||||||||||||
     108 GCAGGCCAGGCCCAGAGGATGGAGAGAACACTGCCCAGTGGAGGAGCCAG 157

      51 GluAsnGluGluAspGlyGluGluAspProAspArgTyrValCysSerGl 67
         |||||||||:::|||   ||||||||||||||:::|||:::||||||||
     158 GAGAACGAGGACGACTGTGAGGAGGACCCGGACCACTATGCCTGCAGTGG 207

      67 yValProGlyArgProProGlyLeuGluGluGluLeuThrLeuLysTyrG 84
         ||||||||||||||   ||||||||||||||||||||||||||||||||
     208 GGTTCCTGGGCGACCATCGGGCCTGGAGGAGGAGCTGACCCTCAAGTATG 257

      84 lyAlaLysHisValIleMetLeuPheValProValThrLeuCysMetIle 100
         |||||:::   ||||||||||||||||||||||||||||||||||||||||
     258 GGGCGAACGACGTGATCATGCTGTTCGTGCCTGTCACGCTGTGCATGATT 307

     101 ValValValAlaThrIleLysSerValArgPheTyrThrGluLysAsnGl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||

     308 GTGGTGGTGGCCACCATCAAGTCCGTGCGTTTCTACACGGAGAAGAACGG 357

     117 yGlnLeuIleTyrThrProPheThrGluAspThrProSerValGlyGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     358 GCAGCTCATCTACACGCCCTTCACCGAGGACACACCCTCTGTGGGCCAGC 407

     134 rgLeuLeuAsnSerValLeuAsnThrLeuIleMetIleSerValIleVal 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     408 GGCTCCTCAACTCCGTGCTGAACACTCTGATCATGATCAGCGTCATTGTG 457

     151 ValMetThrIlePheLeuValValLeuTyrLysTyrArgCysTyrLysPh 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     458 GTCATGACTATCTTCCTAGTGGTACTCTACAAATATCGCTGTTACAAGTT 507

     167 eIleHisGlyTrpLeuIleMetSerSerLeuMetLeuLeuPheLeuPheT 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     508 CATCCATGGCTGGCTGATCATGTCCTCCCTGATGCTCCTCTTCTTGTTCA 557

     184 hrTyrIleTyrLeuGlyGluValLeuLysThrTyrAsnValAlaMetAsp 200
         ||||||||||||||||||||||:::|||||||||||||||||||||||||
     558 CCTACATCTACCTCGGGGAAGTGTTCAAGACCTACAATGTGGCCATGGAC 607

     201 TyrProThrLeuLeuLeuThrValTrpAsnPheGlyAlaValGlyMetVa 217
         ||||||||||||:::|||:::|||||||||||||||||||||||||||||
     608 TACCCCACACTGTTCCTGGCTGTCTGGAACTTCGGGGCAGTGGGCATGGT 657

     217 lCysIleHisTrpLysGlyProLeuValLeuGlnGlnAlaTyrLeuIleM 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||:
     658 GTGCATCCACTGGAAGGGTCCCCTGGTGCTCCAGCAGGCTTACCTCATTG 707

     234 etIleSerAlaLeuMetAlaLeuValPheIleLysTyrLeuProGluTrp 250
         ::||||||||||||||||||||||||||||||||||||||||||||||||
     708 TGATCAGCGCACTCATGGCCCTGGTGTTCATCAAGTACCTGCCGGAGTGG 757

     251 SerAlaTrpValIleLeuGlyAlaIleSerValTyrAspLeuValAlaVa 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     758 TCCGCTTGGGTCATCTTGGGCGCCATCTCTGTGTACGATCTCGTGGCTGT 807

     267 lLeuCysProLysGlyProLeuArgMetLeuValGluThrAlaGlnGluA 284
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     808 GCTGTGCCCCAAAGGGCCACTGAGGATGCTGGTGGAAACTGCCCAGGAAA 857

     284 rgAsnGluProIlePheProAlaLeuIleTyrSerSerAlaMetValTrp 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     858 GAAACGAGCCCATATTTCCTGCCCTCATATACTCATCTGCCATGGTGTGG 907

     301 ThrValGlyMetAlaLysLeuAspProSerSerGlnGlyAlaLeuGlnLe 317
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     908 ACGGTAGGCATGGCGAAGCTGGACCCCTCCTCTCAGGGAGCACTGCAGCT 957

     317 uProTyrAspProGluMetGluGluAspSerTyrAspSerPheGlyGluP 334
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     958 CCCTTACGATCCAGAGATGGAAGAAGACTCCTATGACAGTTTTGGGGAAC 1007

     334 roSerTyrProGluValPheGluProProLeuThrGlyTyrProGlyGlu 350
         ||||||||||||||:::||||||   |||      ||||||||||||||||
    1008 CCTCATACCCTGAAGCCTTCGAAGCTCCCCAGCCTGGCTACCCAGGGGAA 1057

     351 GluLeuGluGluGluGluGluArgGlyValLysLeuGlyLeuGlyAspPh 367
         |||   ||||||||||||||||||||||||||||||||||||||||||||
    1058 GAGCCGGAGGAGGAGGAGGAAAGGGGTGTAAAGCTTGGTCTGGGAGACTT 1107

     367 eIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySerGlyA 384
         ||||||||||||||||||||||||||||||||||||||||||:::|||||
    1108 CATCTTCTACAGCGTTCTGGTGGGCAAGGCAGCAGCCACAGGCAACGGAG 1157

     384 spTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLeuCys 400
         |||||:::|||||||||||||||||:::|||||||||||||||||||||
    1158 ACTGGAGCACGACGCTGGCCTGCTTTATCGCCATCCTCATTGGCTTGTGT 1207

     401 LeuThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAlaLeuPr 417
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1208 CTCACCCTCCTGCTGCTCGCCGTGTTCAAGAAGGCCCTGCCCGCCCTCCC 1257
```

```
     939 TACTCCTCTGCCATGATGTGGACAGTGGGAATGGCAGAT...TCCGCTAC 985

     310 rSerGlnGlyAlaLeu......GlnLeuProTyr.............AspP 321
         :::::::|||    :::      |||:::   :::             |
     986 AGCAGATGGAAGAATGAATCAGCAAGTGCAGCATATAGACAGAAACACAC 1035

     321 roGluMetGluGluAspSerTyrAspSerPheGlyGluProSerTyrPro 337
         |||||      :::::::::   ::::::   :::|||
    1036 CTGAAGGTGCAAACAGCACAGTGGAGGACGCTGCAGAGACCAGGATTCAA 1085

     338 GluValPheGluProProLeuThrGlyTyrProGlyGluGluLeuGluGl 354
                           |||             ::::::   :::
    1086 ....................ACACAAAGTAATCTCTCTTCTGAGGATCC 1114

     354 uGluGluGluArgGlyValLysLeuGlyLeuGlyAspPheIlePheTyrS 371
         :::||||||||||||||||||||||||||||||||||||||||||||||||
    1115 AGATGAGGAAAGAGGAGTGAAGCTTGGTCTTGGAGACTTCATATTCTACA 1164

     371 erValLeuValGlyLysAlaAlaAlaThrGlySerGlyAspTrpAsnThr 387
         ||||||||||||||||||||||||||:::||||||||||||||||||||
    1165 GTGTGCTAGTTGGAAAAGCAGCTGCAACTGCAAGCGGAGACTGGAATACC 1214

     388 ThrLeuAlaCysPheValAlaIleLeuIleGlyLeuCysLeuThrLeuLe 404
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1215 ACATTGGCTTGTTTTGTGGCCATTTTAATAGGTTTATGTCTGACGCTGTT 1264

     404 uLeuLeuAlaValPheLysLysAlaLeuProAlaLeuProIleSerIleT 421
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1265 GTTGCTGGCAGTTTTCAAGAAAGCTCTCCCTGCCCTACCGATTTCTATCA 1314

     421 hrPheGlyLeuIlePheTyrPheSerThrAspAsnLeuValArgProPhe 437
         ||||||||||||||||||||||||||||||:::||||||||||||
    1315 CCTTTGGTTTGATCTTCTACTTTTCTACAGATAATATAGTACGCCCTTTC 1364

     438 MetAspThrLeuAlaSerHisGlnLeuTyrIle 448
         |||||||||||||||||||||||:::||||||
    1365 ATGGACACCCTTGCATCCCACCAAATGTACATT 1397

seq_name: gb_pat:A63557

seq_documentation_block:
LOCUS       A63557       2681 bp    DNA             PAT       12-MAR-1998
DEFINITION  Sequence 10 from Patent WO9722695.
ACCESSION   A63557
VERSION     A63557.1  GI:3717212
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2681)
  AUTHORS   Talerman,A., Amson,R. and Cohen,D.
  TITLE     NUCLEOTIDE SEQUENCES, PROTEINS, DRUGS AND DIAGNOSTIC AGENTS FOR
            TREATING CANCER
  JOURNAL   Patent: WO 9722695-A 10 26-JUN-1997;
            FONDATION JEAN DAUSSET CEPH (FR)
COMMENT     Other publication FR 2747691 19971024
            Other publication FR 2742766 19970627.
FEATURES             Location/Qualifiers
     source          1. .2681
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      663 a    709 c    658 g    650 t      1 others
ORIGIN


alignment_scores:
           Quality: 1468.00             Length:    467
             Ratio:  3.915               Gaps:      6
 Percent Similarity:  80.300   Percent Identity: 64.026

alignment_block:
 US-08-509-359B-137 x A63557   ..
```

```
Align seg 1/1  to: A63557  from: 1  to: 2681

      24 GluSerProThrProArgSerCysGlnGluGlyArgGlnGlyProGluAs 40
         |||   |||:::|||   |||      ::::::   |||      |||||
     261 GAGATACCTGCACCTTTGTCCTACTTCCAGAATGCCCAGATGTCTGAGGA 310

      40 pGlyGluAsnThrAlaGlnTrpArgSerGlnGluAsnGluGluAspGlyG 57
         |:::::::::::::::      ||||||||:::::::::::::::   :
     311 CAGCCACTCCAGCAGCGCCATCCGGAGCCAGAATGACAGCCAAGAACGGC 360

      57 luGluAspProAspArgTyrValCysSerGlyValPro..... 69
         ::::::::   ||||||       ::::::
     361 AGCAGCAGCATGACAGGCAGAGACTTGACAACCCTGAGCCAATATCTAAT 410

      70 GlyArgProProGly.....Le 75
         |||||||||   :::
     411 GGGCGGCCCCAGAGTAACTCAAGACAGGTGGTGGAACAAGATGAGGAGGA 460

      75 uGluGluGluLeuThrLeuLysTyrGlyAlaLysHisValIleMetLeuP 92
         :::||||||||||||||||||||||||||||||||||||||||||||||||
     461 AGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCATGCTCT 510

      92 heValProValThrLeuCysMetIleValValValAlaThrIleLysSer 108
         ||||||||||||||||||||||:::||||||||||||||||||||||||||
     511 TTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCA 560

     109 ValArgPheTyrThrGluLysAsnGlyGlnLeuIleTyrThrProPheTh 125
         |||   |||||||||:::|||:::||||||||||||||||||||||||||
     561 GTCAGCTTCTATACCCGGAAGGACGGTCAGCTAATCTACACCCCATTCAC 610

     125 rGluAspThrProSerValGlyGlnArgLeuLeuAsnSerValLeuAsnT 142
         ||||||||||   :::||||||||||||   |||:::|||:::|||||||:
     611 AGAAGACACTGAGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATG 660

     142 hrLeuIleMetIleSerValIleValValMetThrIlePheLeuValVal 158
         ::   ||||||||||||||||||||:::||||||||:::||||||||||
     661 CGGCCATCATGATCAGTGTCATTGTCATTATGACCATCCTCCTGGTGGTC 710

     159 LeuTyrLysTyrArgCysTyrLysPheIleHisGlyTrpLeuIleMetSe 175
         ||||||||||||||||||||||   ||||||:::||||||||:::||
     711 CTGTATAAATACAGGTGCTACAAGGTCATCCACGCCTGGCTTATTATTTC 760

     175 rSerLeuMetLeuLeuPheLeuPheThrTyrIleTyrLeuGlyGluValL 192
         |||||||:::||||||||:::|||:::::::||||||||||||||||||:
     761 ATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACTTAGGGGAAGTAT 810

     192 euLysThrTyrAsnValAlaMetAspTyrProThrLeuLeuLeuThrVal 208
         ::||||||||||||||||:::|||||||   |||:::   |||   :::
     811 TTAAGACCTACAATGTCGCCGTGGACTACGTTACAGTAGCACTCCTAATC 860

     209 TrpAsnPheGlyAlaValGlyMetValCysIleHisTrpLysGlyProLe 225
         ||||||||||||:::||||||||:::::::||||||||||||||||||||
     861 TGGAATTTTGGTGTGGTCGGGATGATTGCCATCCACTGGAAAGGCCCCCT 910

     225 uValLeuGlnGlnAlaTyrLeuIleMetIleSerAlaLeuMetAlaLeuV 242
         |   ||||||||||||||||||||||||||||||||||||||||||||||||
     911 TCGACTGCAGCAGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGG 960

     242 alPheIleLysTyrLeuProGluTrpSerAlaTrpValIleLeuGlyAla 258
         ||||||||||||||||||||||||||||:::||||||:::||||||:::::
     961 TATTTATCAAGTACCTCCCCGAATGGACCGCATGGCTCATCTTGGCTGTG 1010

     259 IleSerValTyrAspLeuValAlaValLeuCysProLysGlyProLeuAr 275
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1011 ATTTCAGTATATGATTTGGTGGCTGTTTTATGTCCCAAAGGCCCACTTCG 1060

     275 gMetLeuValGluThrAlaGlnGluArgAsnGluProIlePheProAlaL 292
         ||||||||||||||||||||||||||||||||   ::::||||||||||
    1061 TATGCTGGTTGAAACAGCTCAGGAAAGAAATGAGACTCTCTTTCCAGCTC 1110

     292 euIleTyrSerSerAlaMetValTrpThrValGlyMetAlaLysLeuAsp 308
         ||||||||||||:::|||||||||   |||:::||||||:::   |||
    1111 TTATCTATTCCTCAACAATGGTGTGGTTGGTGAATATGGCTGAAGGAGAC 1160
```

```
     |||||||||||||:::|||||||||   |||:::||||||:::   |||
1042 TTATCTATTCCTCAACAATGGTGTGGTTGGTGAATATGGCTGAAGGAGAC 1091

 309 ProSerSerGlnGlyAlaLeuGlnLeuProTyrAspProGluMetGlu.. 324
     |||:::::::|||         ::::::|||   :::|||:::   :::
1092 CCAGAAGCCCAA......AGGAGGGTACCCAAGAACCCCAAGTATAACAC 1135

 325 .....GluAspSerTyrAspSerPheGlyGluP 334
          :::||||||   ::::::      :::
1136 ACAAAGAGCGGAGAGAGAGACACAGGACAGTGGTTCTGGGAACGATGATG 1185

 334 roSerTyrProGluValPheGluProProLeuThrGlyTyrProGly... 349
       :::::::   |||   :::|||            ::::::   |||
1186 GTGGCTTCAGTGAGGAGTGGGAGGCCCAAAGAGACAGTCACCTGGGGCCT 1235

 350 .............................GluGluLeu..... 352
                                 :::||||||
1236 CATCGCTCCACTCCCGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGAG 1285

 353 ............GluGluGluGluGluArgGlyValLysLeuGlyLeuG 365
                 |||:::   ||||||||||||||||||||||||||||
1286 CATTCTAACGAGTGAAGACCCGGAGGAAAGAGGAGTAAAACTTGGACTGG 1335

 365 lyAspPheIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGly 381
     ||||||||||||||||||||||||||||||||||||:::|||||||:::
1336 GAGATTTCATTTTCTACAGTGTTCTGGTTGGTAAGGCCTCAGCAACCGCC 1385

 382 SerGlyAspTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGl 398
     |||||||||||||||||||:::||||||||||||||||||||||||||
1386 AGTGGAGACTGGAACACAACCATAGCCTGCTTTGTAGCCATACTGATCGG 1435

 398 yLeuCysLeuThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProA 415
     ||||||||||||||||||||||||||||:::||||||||||||||||||
1436 CCTGTGCCTTACATTACTCCTGCTCGCCATTTTCAAGAAAGCGTTGCCAG 1485

 415 laLeuProIleSerIleThrPheGlyLeuIlePheTyrPheSerThrAsp 431
     ||||||||||||||||||||||||||||||:::|||||||||:::||||||
1486 CCCTCCCCATCTCCATCACCTTCGGGCTCGTGTTCTACTTCGCCACGGAT 1535

 432 AsnLeuValArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIl 448
        ||||||:::||||||||||||   ||||||   ||||||:::|||||
1536 TACCTTGTGCAGCCCTTCATGGACCAACTTGCATTCCATCAGTTTTATAT 1585

 448 e 448
     |
1586 C 1586
```

```
seq_name: gb_ro:MUSS1PR

seq_documentation_block:
LOCUS       MUSS1PR      1964 bp    mRNA            ROD       24-JUL-1995
DEFINITION  Mus musculus S182 protein mRNA, complete cds.
ACCESSION   L42177
VERSION     L42177.1  GI:904129
KEYWORDS    S182 protein.
SOURCE      Mus musculus (clone: 917-935) adult brain cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1964)
  AUTHORS   Sherrington,R., Rogaev,E.I., Liang,Y., Rogaeva,E.A., Levesque,G.,
            Ikeda,M., Chi,H., Lin,C., Li,G., Holman,K., Tsuda,T., Mar,L.,
            Foncin,J.-F., Bruni,A.C., Montesi,M.P., Sorbl,S., Rainero,I.,
            Pinessi,L., Nee,L., Chumakov,I., Pollen,D., Brookes,A., Sanseau,P.,
            Polinsky,R.J., Wasco,W., Da Silva,H.A.R., Haines,J.L.,
            Pericak-Vance,M.A., Tanzi,R.E., Roses,A.D., Fraser,P.E.,
            Rommens,J.M. and St. George-Hyslop,P.H.
  TITLE     Cloning of a gene bearing missense mutations in early-onset
            familial Alzheimer's disease
  JOURNAL   Nature 375 (6534), 754-760 (1995)
  MEDLINE   95319502
COMMENT     On Jul 25, 1995 this sequence version replaced gi:897616.
FEATURES             Location/Qualifiers
     source          1. .1964
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="917-935"
                     /dev_stage="adult"
                     /germline
                     /tissue_type="brain"
                     /map="14q24.3"
                     /note="clone:  composite sequence from a ~ 0.7 kb RT-PCR
                     clone and an overlapping ~2.0 kb cDNA clone mc6-3-2
                     recovered from adult mouse brain cDNA library in
                     Bluescript at the XhoI site."
     5'UTR           1. .187
     mRNA            1. .1964
     CDS             188. .1591
                     /note="murine homologue of human S182 gene (accession
                     number L42110); putative"
                     /codon_start=1
                     /function="unknown"
                     /product="S182 protein"
                     /protein_id="AAC42094.1"
                     /db_xref="GI:904130"
                     /translation="MTEIPAPLSYFQNAQMSEDSHSSSAIRSQNDSQERQQQHDRQRL
                     DNPEPISNGRPQSNSRQVVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKS
                     VSFYTRKDGQLIYTPFTEDTETVGQRALHSILNAAIMISVIVIMTILLVVLYKYRCYK
                     VIHAWLIISSLLLLFFFSFIYLGEVFKTYNVAVDYVTVALLIWNFGVVGMIAIHWKGP
                     LRLQQAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQ
                     ERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVPKNPKYNTQRAERETQDSGSGNDDG
                     GFSEEWEAQRDSHLGPHRSTPESRAAVQELSGSILTSEDPEERGVKLGLGDFIFYSVL
                     VGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYFATD
                     YLVQPFMDQLAFHQFYI"
     3'UTR           1592. .1964
BASE COUNT      503 a    503 c    496 g    460 t      2 others
ORIGIN


alignment_scores:
            Quality: 1468.00             Length:    467
              Ratio:  3.915               Gaps:      6
 Percent Similarity:  80.300   Percent Identity: 64.026

alignment_block:
 US-08-509-359B-137 x MUSS1PR   ..

 Align seg 1/1  to: MUSS1PR  from: 1  to: 1964

  24 GluSerProThrProArgSerCysGlnGluGlyArgGlnGlyProGluAs 40
     |||   |||:::|||   |||      ::::::   |||      |||||
 194 GAGATACCTGCACCTTTGTCCTACTTCCAGAATGCCCAGATGTCTGAGGA 243

  40 pGlyGluAsnThrAlaGlnTrpArgSerGlnGluAsnGluGluAspGlyG 57
     |:::::::::::::::      ||||||||:::::::::::::::   :
 244 CAGCCACTCCAGCAGCGCCATCCGGAGCCAGAATGACAGCCAAGAACGGC 293

  57 luGluAspProAspArgTyrValCysSerGlyValPro..... 69
     ::::::::   ||||||        ::::::
 294 AGCAGCAGCATGACAGGCAGAGACTTGACAACCCTGAGCCAATATCTAAT 343

  70 GlyArgProProGly.............................Le 75
     |||||||||   :::
 344 GGGCGGCCCCAGAGTAACTCAAGACAGGTGGTGGAACAAGATGAGGAGGA 393

  75 uGluGluGluLeuThrLeuLysTyrGlyAlaLysHisValIleMetLeuP 92
      :::||||||||||||||||||||||||||||||||||||||||||||||
 394 AGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCATGCTCT 443

  92 heValProValThrLeuCysMetIleValValValAlaThrIleLysSer 108
     ||||||||||||||||||||||:::|||||||||||||||||||||||||
 444 TTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCA 493

 109 ValArgPheTyrThrGluLysAsnGlyGlnLeuIleTyrThrProPheTh 125
     |||   |||||||||:::|||:::|||||||||||||||||||||||||
 494 GTCAGCTTCTATACCCGGAAGGACGGTCAGCTAATCTACACCCCATTCAC 543
```

```
OM of: US-08-509-359B-137 to: N_Geneseq_36:*   out_format : pfs

Date: Mar 18, 2000 11:46 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-1998 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US08509359/runat_17032000_164056_19497/app_query.fasta.1
 -DB=N_Geneseq_36 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=1000000 -USER=US08509359 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT
 -THREADS=1

Search information block:
Query: US-08-509-359B-137
Query length: 448
Database: N_Geneseq_36:*
Database sequences: 311585
Database length: 125096042
Search time (sec): 37.450000



score_list:
Sequence           Strd Orig    ZScore    EScore Len  ! Documentation  ..
N_Geneseq_36:T40031  + 2336.00 4306.70  2.8e-232  2229  ! Human presenilin-2 wild type c
N_Geneseq_36:T51253  + 2336.00 4306.67  2.8e-232  2236  ! Human AD4 protein coding seque
N_Geneseq_36:V04669  + 2336.00 4306.70  2.8e-232  2229  ! Human presenilin-2 cDNA (hPS2)
N_Geneseq_36:T87426  + 2320.50 4277.65  1.1e-230  2276  ! Full AD4/AD3LP sequence. Ident
N_Geneseq_36:T87401  + 1623.50 2987.65  8.2e-159  1417  ! AD4/AD3LP sequence. Identifyin
N_Geneseq_36:T40030  + 1468.00 2694.82  1.7e-142  1964  ! Murine presenilin-1 wild type
N_Geneseq_36:T64819  + 1468.00 2691.19  2.7e-142  2681  ! Tumour suppressor inhibited pa
N_Geneseq_36:V04668  + 1468.00 2694.82  1.7e-142  1964  ! Mouse presenilin-1 cDNA. New i
N_Geneseq_36:T40028  + 1467.00 2688.97  3.5e-142  2765  ! Presenilin-1-1 wild type codin
N_Geneseq_36:V17358  + 1467.00 2688.98  3.5e-142  2764  ! PS1/467 protein coding sequenc
N_Geneseq_36:V04666  + 1467.00 2688.97  3.5e-142  2765  ! Human presenilin-1 cDNA (hPS1-
N_Geneseq_36:T85332  + 1462.00 2679.68  1.2e-141  2765  ! Human S182 gene, PS1 locus, re
N_Geneseq_36:V29525  + 1460.00 2675.97  1.9e-141  2764  ! Homo sapiens PS-1 cDNA. Nuclei
N_Geneseq_36:T59535  + 1454.00 2670.06  4.0e-141  1762  ! Human early onset Alzheimer's
N_Geneseq_36:T59536  + 1454.00 2670.14  4.0e-141  1750  ! Human early onset Alzheimer's
N_Geneseq_36:T75576  + 1449.00 2659.80  1.5e-140  1914  ! Presenilin-1 VRSQ variant gene
N_Geneseq_36:T40029  + 1448.00 2652.38  3.9e-140  3086  ! Presenilin-1-2, alternatively
N_Geneseq_36:V04667  + 1448.00 2652.38  3.9e-140  3086  ! Human presenilin-1 cDNA (hPS1-
N_Geneseq_36:V03246  + 1446.00 2655.27  2.7e-140  1750  ! Human presenilin I cDNA. DNA e
N_Geneseq_36:T63207  + 1442.00 2646.81  7.9e-140  1911  ! Human S182 gene associated wit
N_Geneseq_36:T87402  + 1438.00 2642.29  1.4e-139  1488  ! Partial AD3 sequence. Identify
N_Geneseq_36:V17357  + 1438.00 2640.72  1.7e-139  1703  ! PS1/429 protein coding sequenc
N_Geneseq_36:T85333  + 1435.00 2629.50  7.3e-139  2765  ! Human mutant S182 gene, PS1 lo
N_Geneseq_36:T40043  + 1166.50 2134.83  2.6e-111  1895  ! Presenilin homologue. New pres
N_Geneseq_36:T60306  + 1003.50 1834.59  1.4e-94   1500  ! Caenorhabditis elegans SEL-12
N_Geneseq_36:T51271  +  666.00 1220.72  2.1e-60   473  ! Human expressed sequence tag ES
N_Geneseq_36:T51258  +  385.00  679.56  3.0e-30   2387  ! Human AD4 gene genomic sequenc
N_Geneseq_36:T51256  +  377.00  660.75  3.3e-29   3349  ! Human AD4 gene genomic sequenc
N_Geneseq_36:X12881  +  348.00  638.06  6.1e-28   230  ! Human biallelic polymorphic DNA
N_Geneseq_36:X11761  +  348.00  638.06  6.1e-28   230  ! Human biallelic polymorphic DNA
N_Geneseq_36:T40037  +  315.00  560.25  1.3e-23   945  ! Presenilin-1 exon 8. New presen
N_Geneseq_36:T99666  +  315.00  560.25  1.3e-23   945  ! Human presenilin-1 gene exon 8
N_Geneseq_36:T51260  +  270.00  459.79  5.2e-18   4004  ! Human AD4 gene genomic sequenc
N_Geneseq_36:T51259  +  262.00  452.67  1.3e-17   2058  ! Human AD4 gene genomic sequenc
N_Geneseq_36:T51257  +  229.50  396.44  1.8e-14   1438  ! Human AD4 gene genomic sequenc
N_Geneseq_36:T40042  +  224.00  394.02  2.4e-14   736  ! Presenilin-1 exon 13. New prese
N_Geneseq_36:T99671  +  224.00  394.02  2.4e-14   736  ! Human presenilin-1 gene exon 13
N_Geneseq_36:T40034  +  212.50  362.71  1.3e-12   1727  ! Presenilin-1 exon 5. New prese
N_Geneseq_36:T99663  +  212.50  362.71  1.3e-12   1727  ! Human presenilin-1 gene exon 5
N_Geneseq_36:T40035  +  201.00  340.33  2.3e-11   1883  ! Presenilin-1 exon 6. New prese
N_Geneseq_36:T99664  +  201.00  340.33  2.3e-11   1883  ! Human presenilin-1 gene exon 6
N_Geneseq_36:T40041  +  191.00  329.08  9.9e-11   1003  ! Presenilin-1 exon 12. New p
N_Geneseq_36:T99670  +  191.00  329.08  9.9e-11   1003  ! Human presenilin-1 gene exo
N_Geneseq_36:T59473  +  186.00  319.87  3.2e-10   996  ! Early onset Alzheimer's dise
N_Geneseq_36:T40038  +  167.50  292.61  1.1e-08   540  ! Presenilin-1 exon 9. New pre

seq_name: N_Geneseq_36:T40031

seq_documentation_block:
ID   T40031 standard; DNA; 2229 BP.
AC   T40031;
DT   25-JUL-1997  (first entry)
DE   Human presenilin-2 wild type coding sequence.
KW   Presenilin-2; human; hPS1-1; hPS1-2; PS-2; integral membrane protein; AD;
KW   familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW   depression; antibody; gene expression modulator; therapy; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   cds             366. .1712
FT                   /*tag= a
FT                   /product= presenilin-2
PN   WO9634099-A2.
PD   31-OCT-1996.
PF   29-APR-1996; CA0263.
PR   28-APR-1995; US-431048.
PR   28-JUN-1995; US-496841.
PR   31-JUL-1995; US-509359.
PA   (HSCR-) HSC RES & DEV LP.
PA   (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
PI   Fraser PE,  Rommens JM,  St George-Hyslop PH;
DR   WPI; 96-497631/49.
DR   P-PSDB; W05762.
PT   New presenilin genes - useful for diagnosis, therapy and drug
PT   screening of familial Alzheimer's disease, cerebral disorders, etc.
PS   Claim 8; Page 148-150; 178pp; English.
CC   This sequence represents the coding sequence for the human presenilin-2.
CC   T40028 and T40029 represent the coding sequences for the two different
CC   forms of wild type human presenilin-1 (PS-1). The form represented by
CC   T40029 results from alternate splicing of the genomic DNA sequence.
CC   T40030 represents the coding sequence for wild type mouse PS-1. The
CC   presenilins are a family of highly conserved integral membrane proteins
CC   with a common structural motif, common alternate splicing patterns, and
CC   common mutational hot spot regions. Mutations in PS genes are implicated
CC   in familial Alzheimer's disease (AD) and possibly other diseases such as
CC   cerebral haemorrhage, schizophrenia, depression etc., so detection of
CC   mutations in these sequences can be used for diagnosis of these diseases.
CC   The encoded proteins, or vectors that express them or containing
CC   antisense sequences, antibodies selective for mutant forms of the encoded
CC   proteins (such as W05736) and modulators of PS gene expression are
CC   potentially useful for treatment of AD etc. Transgenic animals are useful
CC   as models for drug screening. The antibodies can also be used e.g. for
CC   affinity purification and in immunoassays.
SQ   Sequence   2229 BP;    481 A;    579 C;    633 G;    521 T;


alignment_scores:
            Quality: 2336.00              Length:    448
              Ratio:  5.214                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-509-359B-137 x T40031   ..

 Align seg 1/1  to: T40031  from: 1  to: 2229

       1 MetLeuThrPheMetAlaSerAspSerGluGluGluValCysAspGluAr 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     366 ATGCTCACATTCATGGCCTCTGACAGCGAGGAAGAAGTGTGTGATGAGCG 415

      17 gThrSerLeuMetSerAlaGluSerProThrProArgSerCysGlnGluG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     416 GACGTCCCTAATGTCGGCCGAGAGCCCCACGCCGCGCTCCTGCCAGGAGG 465

      34 lyArgGlnGlyProGluAspGlyGluAsnThrAlaGlnTrpArgSerGln 50
```

```
CC   acid probes specific for the mutant gene, provides a means of
CC   diagnosing Alzheimer's disease.
SQ   Sequence   2236 BP;    488 A;    584 C;    645 G;    519 T;


alignment_scores:
            Quality: 2336.00              Length:    448
              Ratio:  5.214                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-509-359B-137 x T51253   ..

 Align seg 1/1  to: T51253  from: 1  to: 2236

       1 MetLeuThrPheMetAlaSerAspSerGluGluGluValCysAspGluAr 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     368 ATGCTCACATTCATGGCCTCTGACAGCGAGGAAGAAGTGTGTGATGAGCG 417

      17 gThrSerLeuMetSerAlaGluSerProThrProArgSerCysGlnGluG 34
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     418 GACGTCCCTAATGTCGGCCGAGAGCCCCACGCCGCGCTCCTGCCAGGAGG 467

      34 lyArgGlnGlyProGluAspGlyGluAsnThrAlaGlnTrpArgSerGln 50
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     468 GCAGGCAGGGCCCAGAGGATGGAGAGAACACTGCCCAGTGGAGAAGCCAG 517

      51 GluAsnGluGluAspGlyGluGluAspProAspArgTyrValCysSerGl 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     518 GAGAACGAGGAGGACGGTGAGGAGGACCCTGACCGCTATGTCTGTAGTGG 567

      67 yValProGlyArgProProGlyLeuGluGluGluLeuThrLeuLysTyrG 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     568 GGTTCCCGGGCGGCCGCCAGGCCTGGAGGAAGAGCTGACCCTCAAATACG 617

      84 lyAlaLysHisValIleMetLeuPheValProValThrLeuCysMetIle 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     618 GAGCGAAGCACGTGATCATGCTGTTTGTGCCTGTCACTCTGTGCATGATC 667

     101 ValValValAlaThrIleLysSerValArgPheTyrThrGluLysAsnGl 117
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     668 GTGGTGGTAGCCACCATCAAGTCTGTGCGCTTCTACACAGAGAAGAATGG 717

     117 yGlnLeuIleTyrThrProPheThrGluAspThrProSerValGlyGlnA 134
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     718 ACAGCTCATCTACACGCCATTCACTGAGGACACACCCTCGGTGGGCCAGC 767

     134 rgLeuLeuAsnSerValLeuAsnThrLeuIleMetIleSerValIleVal 150
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     768 GCCTCCTCAACTCCGTGCTGAACACCCTCATCATGATCAGCGTCATCGTG 817

     151 ValMetThrIlePheLeuValValLeuTyrLysTyrArgCysTyrLysPh 167
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     818 GTTATGACCATCTTCTTGGTGGTGCTCTACAAGTACCGCTGCTACAAGTT 867

     167 eIleHisGlyTrpLeuIleMetSerSerLeuMetLeuLeuPheLeuPheT 184
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     868 CATCCATGGCTGGTTGATCATGTCTTCACTGATGCTGCTGTTCCTCTTCA 917

     184 hrTyrIleTyrLeuGlyGluValLeuLysThrTyrAsnValAlaMetAsp 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     918 CCTATATCTACCTTGGGGAAGTGCTCAAGACCTACAATGTGGCCATGGAC 967

     201 TyrProThrLeuLeuLeuThrValTrpAsnPheGlyAlaValGlyMetVa 217
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     968 TACCCCACCCTCTTGCTGACTGTCTGGAACTTCGGGGCAGTGGGCATGGT 1017

     217 lCysIleHisTrpLysGlyProLeuValLeuGlnGlnAlaTyrLeuIleM 234
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1018 GTGCATCCACTGGAAGGGCCCTCTGGTGCTGCAGCAGGCCTACCTCATCA 1067

     234 etIleSerAlaLeuMetAlaLeuValPheIleLysTyrLeuProGluTrp 250
         |||||||||||||||||||||||||||||||||||||||||||||||||||

    1068 TGATCAGTGCGCTCATGGCCCTAGTGTTCATCAAGTACCTCCCAGAGTGG 1117

     251 SerAlaTrpValIleLeuGlyAlaIleSerValTyrAspLeuValAlaVa 267
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1118 TCCGCGTGGGTCATCCTGGGCGCCATCTCTGTGTATGATCTCGTGGCTGT 1167

     267 lLeuCysProLysGlyProLeuArgMetLeuValGluThrAlaGlnGluA 284
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1168 GCTGTGTCCCAAAGGGCCTCTGAGAATGCTGGTAGAAACTGCCCAGGAGA 1217

     284 rgAsnGluProIlePheProAlaLeuIleTyrSerSerAlaMetValTrp 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1218 GAAATGAGCCCATATTCCCTGCCCTGATATACTCATCTGCCATGGTGTGG 1267

     301 ThrValGlyMetAlaLysLeuAspProSerSerGlnGlyAlaLeuGlnLe 317
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1268 ACGGTTGGCATGGCGAAGCTGGACCCCTCCTCTCAGGGTGCCCTCCAGCT 1317

     317 uProTyrAspProGluMetGluGluAspSerTyrAspSerPheGlyGluP 334
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1318 CCCCTACGACCCGGAGATGGAAGAAGACTCCTATGACAGTTTTGGGGAGC 1367

     334 roSerTyrProGluValPheGluProProLeuThrGlyTyrProGlyGlu 350
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1368 CTTCATACCCCGAAGTCTTTGAGCCTCCCTTGACTGGCTACCCAGGGGAG 1417

     351 GluLeuGluGluGluGluGluArgGlyValLysLeuGlyLeuGlyAspPh 367
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1418 GAGCTGGAGGAAGAGGAGGAAAGGGGCGTGAAGCTTGGCCTCGGGGACTT 1467

     367 eIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySerGlyA 384
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1468 CATCTTCTACAGTGTGCTGGTGGGCAAGGCGGCTGCCACGGGCAGCGGGG 1517

     384 spTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLeuCys 400
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1518 ACTGGAATACCACGCTGGCCTGCTTCGTGGCCATCCTCATTGGCTTGTGT 1567

     401 LeuThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAlaLeuPr 417
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1568 CTGACCCTCCTGCTGCTTGCTGTGTTCAAGAAGGCGCTGCCCGCCCTCCC 1617

     417 oIleSerIleThrPheGlyLeuIlePheTyrPheSerThrAspAsnLeuV 434
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1618 CATCTCCATCACGTTCGGGCTCATCTTTTACTTCTCCACGGACAACCTGG 1667

     434 alArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIle 448
         |||||||||||||||||||||||||||||||||||||||||||
    1668 TGCGGCCGTTCATGGACACCCTGGCCTCCCATCAGCTCTACATC 1711

seq_name: N_Geneseq_36:V04669

seq_documentation_block:
ID   V04669 standard; cDNA; 2229 BP.
AC   V04669;
DT   20-JUL-1998  (first entry)
DE   Human presenilin-2 cDNA (hPS2).
KW   Presenilin-1; PS1 gene; human; familial Alzheimer's disease; FAD;
KW   cerebral haemorrhage; schizophrenia; depression; epilepsy;
KW   mental retardation; diagnosis; therapy; transgenic animal; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             366. .1715
FT                   /*tag= a
FT   mutation        787
FT                   /*tag= p
FT                   /note= "A to T FAD mutation site (Asn141Ile)"
FT   mutation        1080
FT                   /*tag= q
FT                   /note= "A to G FAD mutation site (Met239Val)"
FT   mutation        1624
FT                   /*tag= r
FT                   /note= "T to C FAD mutation site (Ile420Thr)"
```

```
     |||||||||||||||||||||||||||||||||||||||||||||||||||
  1616 CATCTCCATCACGTTCGGGCTCATCTTTTACTTCTCCACGGACAACCTGG 1665

   434 alArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIle 448
       ||||||||||||||||||||||||||||||||||||||||||||
  1666 TGCGGCCGTTCATGGACACCCTGGCCTCCCATCAGCTCTACATC 1709

seq_name: N_Geneseq_36:T87426

seq_documentation_block:
ID   T87426 standard; DNA; 2276 BP.
AC   T87426;
DT   07-DEC-1997  (first entry)
DE   Full AD4/AD3LP sequence.
KW   AD3; AD4/AD3LP; Alzheimer's disease; chromosome; missegregation;
KW   presenilin; inhibitor; AD; trisomy 21; ss.
OS   Homo sapiens.
PN   WO9707213-A2.
PD   27-FEB-1997.
PF   15-AUG-1996; U13314.
PR   16-AUG-1995; US-002448.
PA   (HARD ) HARVARD COLLEGE.
PI   Li J,  Potter H;
DR   WPI; 97-165297/15.
DR   P-PSDB; W28508.
PT   Identifying genes which cause chromosome missegregation - useful for
PT   identifying causes of and treatments for diseases, e.g. Alzheimer's
PT   disease, cancer and ageing
PS   Claim 28; Fig 28; 77pp; English.
CC   Identifying genes which cause improper chromosome segregation,
CC   screening for inhibitors of chromosome missegregation and processes
CC   caused by genes encoding chromosome missegregation promoters
CC   was exemplified using Alzheimer's disease.  The sequences
CC   given in T87401 to T87426 can be used in the above methods.
CC   It is not clear from the figure legend, the figure and the
CC   disclosure of the specification which sequence of Fig 1 and Fig 28
CC   is the AD4/AD3LP or the AD3 sequence.
SQ   Sequence   2276 BP;    494 A;    595 C;    662 G;    525 T;


alignment_scores:
            Quality: 2320.50              Length:    448
              Ratio:  5.191                Gaps:      1
 Percent Similarity:  99.777   Percent Identity: 99.777

alignment_block:
 US-08-509-359B-137 x T87426   ..

 Align seg 1/1  to: T87426  from: 1  to: 2276

       1 MetLeuThrPheMetAlaSerAspSerGluGluGluValCysAspGluAr 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     411 ATGCTCACATTCATGGCCTCTGACAGCGAGGAAGAAGTGTGTGATGAGCG 460

      17 gThrSerLeuMetSerAlaGluSerProThrProArgSerCysGlnGluG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     461 GACGTCCCTAATGTCGGCCGAGAGCCCCACGCCGCGCTCCTGCCAGGAGG 510

      34 lyArgGlnGlyProGluAspGlyGluAsnThrAlaGlnTrpArgSerGln 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     511 GCAGGCAGGGCCCAGAGGATGGAGAGAACACTGCCCAGTGGAGAAGCCAG 560

      51 GluAsnGluGluAspGlyGluGluAspProAspArgTyrValCysSerGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     561 GAGAACGAGGAGGACGGTGAGGAGGACCCTGACCGCTATGTCTGTAGTGG 610

      67 yValProGlyArgProProGlyLeuGluGluGluLeuThrLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     611 GGTTCCCGGGCGGCCGCCAGGCCTGGAGGAAGAGCTGACCCTCAAATACG 660

      84 lyAlaLysHisValIleMetLeuPheValProValThrLeuCysMetIle 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     661 GAGCGAAGCACGTGATCATGCTGTTTGTGCCTGTCACTCTGTGCATGATC 710

     101 ValValValAlaThrIleLysSerValArgPheTyrThrGluLysAsnGl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     711 GTGGTGGTAGCCACCATCAAGTCTGTGCGCTTCTACACAGAGAAGAATGG 760

     117 yGlnLeuIleTyrThrProPheThrGluAspThrProSerValGlyGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     761 ACAGCTCATCTACACGCCATTCACTGAGGACACACCCTCGGTGGGCCAGC 810

     134 rgLeuLeuAsnSerValLeuAsnThrLeuIleMetIleSerValIleVal 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     811 GCCTCCTCAACTCCGTGCTGAACACCCTCATCATGATCAGCGTCATCGTG 860

     151 ValMetThrIlePheLeuValValLeuTyrLysTyrArgCysTyrLysPh 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     861 GTTATGACCATCTTCTTGGTGGTGCTCTACAAGTACCGCTGCTACAAGTT 910

     167 eIleHisGlyTrpLeuIleMetSerSerLeuMetLeuLeuPheLeuPheT 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     911 CATCCATGGCTGGTTGATCATGTCTTCACTGATGCTGCTGTTCCTCTTCA 960

     184 hrTyrIleTyrLeuGlyGluValLeuLysThrTyrAsnValAlaMetAsp 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     961 CCTATATCTACCTTGGGGAAGTGCTCAAGACCTACAATGTGGCCATGGAC 1010

     201 TyrProThrLeuLeuLeuThrValTrpAsnPheGlyAlaValGlyMetVa 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1011 TACCCCACCCTCTTGCTGACTGTCTGGAACTTCGGGGCAGTGGGCATGGT 1060

     217 lCysIleHisTrpLysGlyProLeuValLeuGlnGlnAlaTyrLeuIleM 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1061 GTGCATCCACTGGAAGGGCCCTCTGGTGCTGCAGCAGGCCTACCTCATCA 1110

     234 etIleSerAlaLeuMetAlaLeuValPheIleLysTyrLeuProGluTrp 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1111 TGATCAGTGCGCTCATGGCCCTAGTGTTCATCAAGTACCTCCCAGAGTGG 1160

     251 SerAlaTrpValIleLeuGlyAlaIleSerValTyrAspLeuValAlaVa 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1161 TCCGCGTGGGTCATCCTGGGCGCCATCTCTGTGTATGATCTCGTGGCTGT 1210

     267 lLeuCysProLysGlyProLeuArgMetLeuValGluThrAlaGlnGluA 284
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1211 GCTGTGTCCCAAAGGGCCTCTGAGAATGCTGGTAGAAACTGCCCAGGAGA 1260

     284 rgAsnGluProIlePheProAlaLeuIleTyrSerSerAlaMetValTrp 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1261 GAAATGAGCCCATATTCCCTGCCCTGATATACTCATCTGCCATGGTGTGG 1310

     301 ThrValGlyMetAlaLysLeuAspProSerSerGlnGlyAlaLeuGlnLe 317
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1311 ACGGTTGGCATGGCGAAGCTGGACCCCTCCTCTCAGGGTGCCCTCCAGCT 1360

     317 uProTyrAspProGluMetGluGluAspSerTyrAspSerPheGlyGluP 334
         ||||||||||||||||||   ||||||||||||||||||||||||||||
    1361 CCCCTACGACCCGGAGATG...GAAGACTCCTATGACAGTTTTGGGGAGC 1407

     334 roSerTyrProGluValPheGluProProLeuThrGlyTyrProGlyGlu 350
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1408 CTTCATACCCCGAAGTCTTTGAGCCTCCCTTGACTGGCTACCCAGGGGAG 1457

     351 GluLeuGluGluGluGluGluArgGlyValLysLeuGlyLeuGlyAspPh 367
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1458 GAGCTGGAGGAAGAGGAGGAAAGGGGCGTGAAGCTTGGCCTCGGGGACTT 1507

     367 eIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySerGlyA 384
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1508 CATCTTCTACAGTGTGCTGGTGGGCAAGGCGGCTGCCACGGGCAGCGGGG 1557

     384 spTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLeuCys 400
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1558 ACTGGAATACCACGCTGGCCTGCTTCGTGGCCATCCTCATTGGCTTGTGT 1607
```

```
seq_name: N_Geneseq_36:T40030

seq_documentation_block:
ID   T40030 standard; DNA; 1964 BP.
AC   T40030;
DT   23-JUL-1997  (first entry)
DE   Murine presenilin-1 wild type coding sequence.
KW   Presenilin-1; mouse; hPS1-1; hPS1-2; PS-2; integral membrane protein; AD;
KW   familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW   depression; antibody; gene expression modulator; therapy; ss.
OS   Mus musculus.
FH   Key             Location/Qualifiers
FT   cds             188. .1591
FT                   /*tag= a
FT                   /product= presenilin-1
PN   WO9634099-A2.
PD   31-OCT-1996.
PF   29-APR-1996; CA0263.
PR   28-APR-1995; US-431048.
PR   28-JUN-1995; US-496841.
PR   31-JUL-1995; US-509359.
PA   (HSCR-) HSC RES & DEV LP.
PA   (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
PI   Fraser PE,  Rommens JM,  St George-Hyslop PH;
DR   WPI; 96-497631/49.
DR   P-PSDB; W05735.
PT   New presenilin genes - useful for diagnosis, therapy and drug
PT   screening of familial Alzheimer's disease, cerebral disorders, etc.
PS   Claim 8; Page 145-146; 178pp; English.
CC   This sequence represents the coding sequence for the murine presenilin-1.
CC   T40028 and T40029 represent the coding sequences for the two different
CC   forms of wild type human presenilin-1 (PS-1). The form represented by
CC   T40029 results from alternate splicing of the genomic DNA sequence.
CC   T40031 represents the coding sequence for wild type human PS-2. The
CC   presenilins are a family of highly conserved integral membrane proteins
CC   with a common structural motif, common alternate splicing patterns, and
CC   common mutational hot spot regions. Mutations in PS genes are implicated
CC   in familial Alzheimer's disease (AD) and possibly other diseases such as
CC   cerebral haemorrhage, schizophrenia, depression etc., so detection of
CC   mutations in these sequences can be used for diagnosis of these diseases.
CC   The encoded proteins, or vectors that express them or containing
CC   antisense sequences, antibodies selective for mutant forms of the encoded
CC   proteins (such as W05736) and modulators of PS gene expression are
CC   potentially useful for treatment of AD etc. Transgenic animals are useful
CC   as models for drug screening. The antibodies can also be used e.g. for
CC   affinity purification and in immunoassays.
SQ   Sequence   1964 BP;    503 A;    503 C;    496 G;    460 T;


alignment_scores:
            Quality: 1468.00              Length:    467
              Ratio:  3.915                Gaps:      6
 Percent Similarity:  80.300   Percent Identity: 64.026

alignment_block:
 US-08-509-359B-137 x T40030   ..

 Align seg 1/1  to: T40030  from: 1  to: 1964

      24 GluSerProThrProArgSerCysGlnGluGlyArgGlnGlyProGluAs 40
         |||   |||:::|||   |||      ::::::  |||     |||||
     194 GAGATACCTGCACCTTTGTCCTACTTCCAGAATGCCCAGATGTCTGAGGA 243

      40 pGlyGluAsnThrAlaGlnTrpArgSerGlnGluAsnGluGluAspGlyG 57
         |:::::::::::::::      |||||||||:::::::::::::::   :
     244 CAGCCACTCCAGCAGCGCCATCCGGAGCCAGAATGACAGCCAAGAACGGC 293

      57 luGluAspProAspArgTyrValCysSerGlyValPro..........  69
         ::::::::   ||||||         ::::::
     294 AGCAGCAGCATGACAGGCAGAGACTTGACAACCCTGAGCCAATATCTAAT 343

      70 GlyArgProProGly:..............................Le 75
         |||||||||   :::

     344 GGGCGGCCCCAGAGTAACTCAAGACAGGTGGTGGAACAAGATGAGGAGGA 393

      75 uGluGluGluLeuThrLeuLysTyrGlyAlaLysHisValIleMetLeuP 92
          :::||||||||||||||||||||||||||||||||||||||||||||||
     394 AGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCATGCTCT 443

      92 heValProValThrLeuCysMetIleValValValAlaThrIleLysSer 108
         |||||||||||||||||||||||:::|||||||||||||||||||||||
     444 TTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCA 493

     109 ValArgPheTyrThrGluLysAsnGlyGlnLeuIleTyrThrProPheTh 125
         |||   |||||||||:::|||:::||||||||||||||||||||||||||
     494 GTCAGCTTCTATACCCGGAAGGACGGTCAGCTAATCTACACCCCATTCAC 543

     125 rGluAspThrProSerValGlyGlnArgLeuLeuAsnSerValLeuAsnT 142
         ||||||||||   :::|||||||||||||   |||:::|||:::||||||:
     544 AGAAGACACTGAGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATG 593

     142 hrLeuIleMetIleSerValIleValValMetThrIlePheLeuValVal 158
         ::   |||||||||||||||||||:::|||||||||:::|||||||||
     594 CGGCCATCATGATCAGTGTCATTGTCATTATGACCATCCTCCTGGTGGTC 643

     159 LeuTyrLysTyrArgCysTyrLysPheIleHisGlyTrpLeuIleMetSe 175
         ||||||||||||||||||||||    ||||||:::||||||||||:::||
     644 CTGTATAAATACAGGTGCTACAAGGTCATCCACGCCTGGCTTATTATTTC 693

     175 rSerLeuMetLeuLeuPheLeuPheThrTyrIleTyrLeuGlyGluValL 192
         |||||||:::|||||||||:::|||:::::::|||||||||||||||||:
     694 ATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACTTAGGGGAAGTAT 743

     192 euLysThrTyrAsnValAlaMetAspTyrProThrLeuLeuLeuThrVal 208
         ::|||||||||||||||||:::||||||   |||:::   |||   :::
     744 TTAAGACCTACAATGTCGCCGTGGACTACGTTACAGTAGCACTCCTAATC 793

     209 TrpAsnPheGlyAlaValGlyMetValCysIleHisTrpLysGlyProLe 225
         ||||||||||||:::|||||||||:::::|||||||||||||||||||||
     794 TGGAATTTTGGTGTGGTCGGGATGATTGCCATCCACTGGAAAGGCCCCCT 843

     225 uValLeuGlnGlnAlaTyrLeuIleMetIleSerAlaLeuMetAlaLeuV 242
         |   ||||||||||||||||||||||||||||||||||||||||||||||
     844 TCGACTGCAGCAGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGG 893

     242 alPheIleLysTyrLeuProGluTrpSerAlaTrpValIleLeuGlyAla 258
         ||||||||||||||||||||||||:::||||||:::||||||::::::
     894 TATTTATCAAGTACCTCCCCGAATGGACCGCATGGCTCATCTTGGCTGTG 943

     259 IleSerValTyrAspLeuValAlaValLeuCysProLysGlyProLeuAr 275
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     944 ATTTCAGTATATGATTTGGTGGCTGTTTTATGTCCCAAAGGCCCACTTCG 993

     275 gMetLeuValGluThrAlaGlnGluArgAsnGluProIlePheProAlaL 292
         |||||||||||||||||||||||||||||||||   :::||||||||||
    994 TATGCTGGTTGAAACAGCTCAGGAAAGAAATGAGACTCTCTTTCCAGCTC 1043

     292 euIleTyrSerSerAlaMetValTrpThrValGlyMetAlaLysLeuAsp 308
         ||||||||||||||:::|||||||||   |||:::|||||||:::   |||
    1044 TTATCTATTCCTCAACAATGGTGTGGTTGGTGAATATGGCTGAAGGAGAC 1093

     309 ProSerSerGlnGlyAlaLeuGlnLeuProTyrAspProGluMetGlu.. 324
         |||:::::|||         :::::::|||   :::|||:::   :::
    1094 CCAGAAGCCCAA......AGGAGGGTACCCAAGAACCCCAAGTATAACAC 1137

     325 .....................GluAspSerTyrAspSerPheGlyGluP 334
                              :::||||||   ::::::      :::
    1138 ACAAAGAGCGGAGAGAGAGACACAGGACAGTGGTTCTGGGAACGATGATG 1187

     334 roSerTyrProGluValPheGluProProLeuThrGlyTyrProGly... 349
           ::::::   |||   :::|||             ::::::   |||
    1188 GTGGCTTCAGTGAGGAGTGGGAGGCCCAAAGAGACAGTCACCTGGGGCCT 1237

     350 ...............................GluGluLeu........ 352
                                        :::||||||
    1238 CATCGCTCCACTCCCGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGAG 1287
```

```
     292 euIleTyrSerSerAlaMetValTrpThrValGlyMetAlaLysLeuAsp 308
         ||||||||||||||:::|||||||||   |||:::|||||||:::   |||
    1111 TTATCTATTCCTCAACAATGGTGTGGTTGGTGAATATGGCTGAAGGAGAC 1160

     309 ProSerSerGlnGlyAlaLeuGlnLeuProTyrAspProGluMetGlu.. 324
         |||::::::|||         :::::::|||   :::|||:::   :::
    1161 CCAGAAGCCCAA......AGGAGGGTACCCAAGAACCCCAAGTATAACAC 1204

     325 .....GluAspSerTyrAspSerPheGlyGluP 334
                 :::||||||   ::::::      :::
    1205 ACAAAGAGCGGAGAGAGAGACACAGGACAGTGGTTCTGGGAACGATGATG 1254

     334 roSerTyrProGluValPheGluProProLeuThrGlyTyrProGly... 349
           ::::::   |||   :::|||           ::::::   |||
    1255 GTGGCTTCAGTGAGGAGTGGGAGGCCCAAAGAGACAGTCACCTGGGGCCT 1304

     350 .....GluGluLeu..... 352
                 :::||||||
    1305 CATCGCTCCACTCCCGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGAG 1354

     353 .....GluGluGluGluGluArgGlyValLysLeuGlyLeuG 365
                 |||:::   |||||||||||||||||||||||||||||
    1355 CATTCTAACGAGTGAAGACCCGGAGGAAAGAGGAGTAAAACTTGGACTGG 1404

     365 lyAspPheIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGly 381
         |||||||||||||||||||||||||||||||||||:::|||||||:::
    1405 GAGATTTCATTTTCTACAGTGTTCTGGTTGGTAAGGCCTCAGCAACCGCC 1454

     382 SerGlyAspTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGl 398
         |||||||||||||||||||:::|||||||||||||||||||||||||||
    1455 AGTGGAGACTGGAACACAACCATAGCCTGCTTTGTAGCCATACTGATCGG 1504

     398 yLeuCysLeuThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProA 415
         |||||||||||||||||||||||||||||:::|||||||||||||||||||
    1505 CCTGTGCCTTACATTACTCCTGCTCGCCATTTTCAAGAAAGCGTTGCCAG 1554

     415 laLeuProIleSerIleThrPheGlyLeuIlePheTyrPheSerThrAsp 431
         |||||||||||||||||||||||||||||::|||||||||:::||||||
    1555 CCCTCCCCATCTCCATCACCTTCGGGCTCGTGTTCTACTTCGCCACGGAT 1604

     432 AsnLeuValArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIl 448
           ||||||:::|||||||||||||   ||||||   ||||||:::||||||
    1605 TACCTTGTGCAGCCCTTCATGGACCAACTTGCATTCCATCAGTTTTATAT 1654

     448 e 448
         |
    1655 C 1655
```

seq_name: N_Geneseq_36:V04668

seq_documentation_block:

```
ID   V04668 standard; cDNA; 1964 BP.
AC   V04668;
DT   20-JUL-1998  (first entry)
DE   Mouse presenilin-1 cDNA.
KW   Presenilin-1; PS1 gene; mouse; familial Alzheimer's disease; FAD;
KW   cerebral haemorrhage; schizophrenia; depression; epilepsy;
KW   mental retardation; diagnosis; therapy; transgenic animal; ss.
OS   Mus musculus.
FH   Key             Location/Qualifiers
FT   CDS             188. .1561
FT                   /*tag=  a
PN   WO9801549-A2.
PD   15-JAN-1998.
PF   04-JUL-1997; CA0475.
PR   02-JAN-1997; US-034590.
PR   05-JUL-1996; US-021673.
PR   12-JUL-1996; US-021700.
PR   08-NOV-1996; US-029895.
PA   (HSCR-) HSC RES & DEV LP.
PA   (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
PI   Fraser PE, Rommens JM, St George-Hyslop PH;
DR   WPI; 98-286355/25.
DR   P-PSDB; W23966.
PT   New isolated mutant presenilin-1 genes - useful for developing
PT   products for use in detection, diagnosis and therapy of Alzheimer's
PT   disease and for drug screening
PS   Disclosure; Page 197-199; 238pp; English.
CC   This cDNA clone for a murine presenilin-1 (PS1) homologue (see
CC   W23966).  It was isolated from a mouse cDNA library using a DNA
CC   probe from the human PS1 gene (see V04666).  Mutations in the
CC   human PS1 and PS2 genes (see V04666-68) have been linked to the
CC   development in humans of forms of familial Alzheimer's disease
CC   (FAD).  All amino acids that are mutated in analysed FAD
CC   pedigrees were conserved in the murine homologue.  Use of the
CC   nucleic acids and proteins comprising or derived from presenilins
CC   can be made in screening and diagnosing FAD, identifying and
CC   developing therapeutics for treatment of FAD, and in producing cell
CC   lines and transgenic animals useful as models of FAD.
SQ   Sequence   1964 BP;    503 A;    503 C;    496 G;    460 T;
```

alignment_scores:

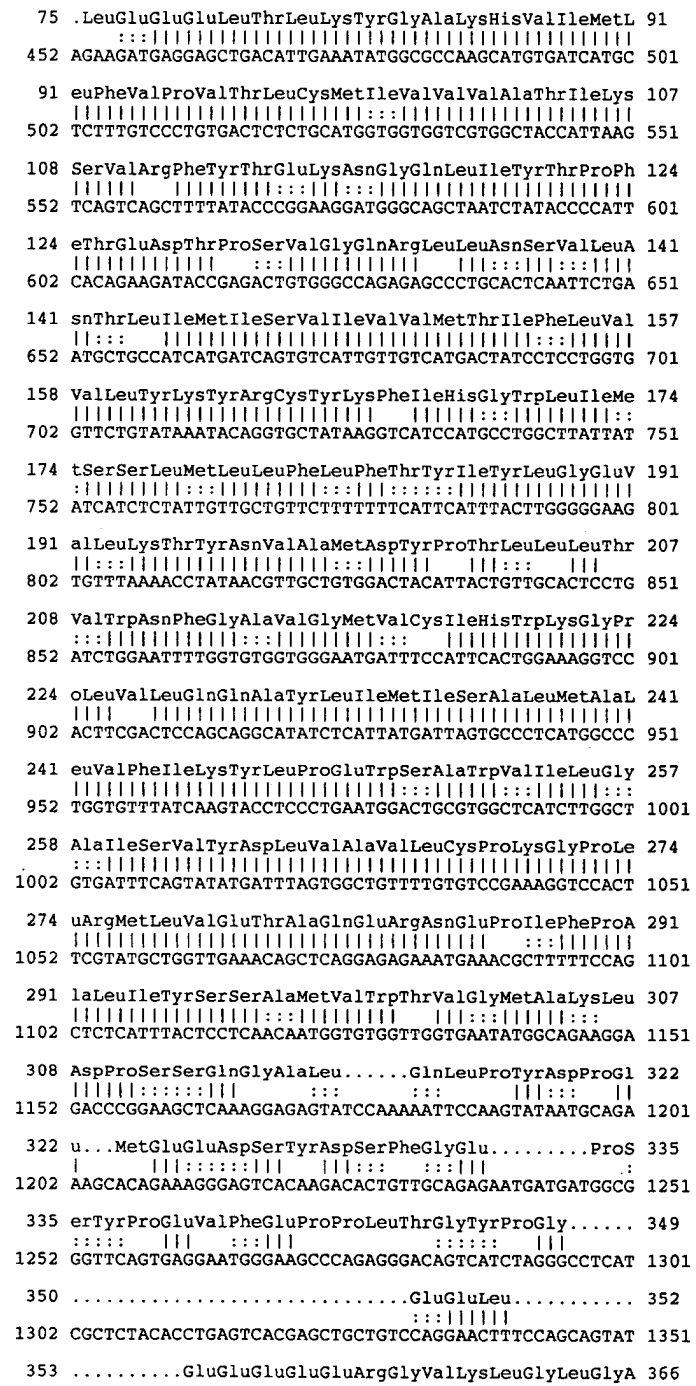
```
            Quality: 1468.00            Length:    467
              Ratio:  3.915              Gaps:      6
 Percent Similarity:  80.300   Percent Identity: 64.026
```

alignment_block:

US-08-509-359B-137 x V04668 ..

Align seg 1/1 to: V04668 from: 1 to: 1964

```
      24 GluSerProThrProArgSerCysGlnGluGlyArgGlnGlyProGluAs 40
         |||   |||:::|||   |||      ::::::   |||      |||||
     194 GAGATACCTGCACCTTTGTCCTACTTCCAGAATGCCCAGATGTCTGAGGA 243

      40 pGlyGluAsnThrAlaGlnTrpArgSerGlnGluAsnGluGluAspGlyG 57
         |:::::::::::::::      |||||||||:::::::::::::::   :
     244 CAGCCACTCCAGCAGCGCCATCCGGAGCCAGAATGACAGCCAAGAACGGC 293

      57 luGluAspProAspArgTyrValCysSerGlyValPro..... 69
         ::::::::   ||||||         ::::::
     294 AGCAGCAGCATGACAGGCAGAGACTTGACAACCCTGAGCCAATATCTAAT 343

      70 GlyArgProProGly.....Le 75
         |||||||||   :::
     344 GGGCGGCCCCAGAGTAACTCAAGACAGGTGGTGGAACAAGATGAGGAGGA 393

      75 uGluGluGluLeuThrLeuLysTyrGlyAlaLysHisValIleMetLeuP 92
          :::||||||||||||||||||||||||||||||||||||||||||||||
     394 AGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCATGCTCT 443

      92 heValProValThrLeuCysMetIleValValValAlaThrIleLysSer 108
         ||||||||||||||||||||||||::|||||||||||||||||||||||
     444 TTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCA 493

     109 ValArgPheTyrThrGluLysAsnGlyGlnLeuIleTyrThrProPheTh 125
         |||   |||||||||:::|||:::|||||||||||||||||||||||||
     494 GTCAGCTTCTATACCCGGAAGGACGGTCAGCTAATCTACACCCCATTCAC 543

     125 rGluAspThrProSerValGlyGlnArgLeuLeuAsnSerValLeuAsnT 142
         ||||||||||   :::||||||||||||   |||:::|||:::||||||:
     544 AGAAGACACTGAGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATG 593

     142 hrLeuIleMetIleSerValIleValValMetThrIlePheLeuValVal 158
         ::   |||||||||||||||||||:::|||||||||:::|||||||||
     594 CGGCCATCATGATCAGTGTCATTGTCATTATGACCATCCTCCTGGTGGTC 643

     159 LeuTyrLysTyrArgCysTyrLysPheIleHisGlyTrpLeuIleMetSe 175
         |||||||||||||||||||||||   ||||||:::|||||||||:::||
     644 CTGTATAAATACAGGTGCTACAAGGTCATCCACGCCTGGCTTATTATTTC 693

     175 rSerLeuMetLeuLeuPheLeuPheThrTyrIleTyrLeuGlyGluValL 192
         |||||||:::|||||||||:::|||:::::::|||||||||||||||||:
     694 ATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACTTAGGGGAAGTAT 743
```



seq_name: N_Geneseq_36:V17358

seq_documentation_block:

ID V17358 standard; DNA; 2764 BP.
AC V17358;
DT 04-JUN-1998 (first entry)
DE PS1/467 protein coding sequence.
KW Presenilin peptide; PS1/429; immunogen; immune response; PS1 gene;
KW Alzheimer's disease; mitochondrial pathology; neurodegeneration;
KW apoptosis; PS1/467; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 249. .1652
FT /*tag= a
PN WO9746678-A1.
PD 11-DEC-1997.
PF 03-JUN-1997; U09272.
PR 18-JUL-1996; US-683315.
PR 06-JUN-1996; US-659296.
PA (FARB) BAYER CORP.
PI Chisholm JC, Davis JN, Drache B;
DR WPI; 98-042186/04.
DR P-PSDB; W41430.
PT DNA encoding presenilin peptide PS1/429 and its analogues - useful
PT for diagnosis and treatment of Alzheimer's disease
PS Disclosure; Fig 2; 77pp; English.
CC This sequence encodes the PS1/467 presenilin peptide. This sequence is
CC specifically stated as not being in the nucleic acid of the invention,
CC which encodes the PS1/429 presenilin peptide PS1/429 (II). Cells
CC transformed with the DNA are used to produce recombinant (II) and
CC analogues, useful e.g. as immunogens for generating an immune response
CC against PS1/429. (II) is a new product of the PS1 gene, mutations in
CC which cause Alzheimer's disease (AD). The nucleic acids are generally
CC useful as probes for detection and quantification of PS1/429,
CC particularly for diagnosis of AD, especially the target sequences that
CC hybridise with probes are isolated for sequencing. Antibodies (Ab) can
CC also be diagnosed at the protein level using Ab as immunoassay reagents.
CC Ab can also be used to identify epitopes and for affinity purification of
CC peptides. Antisense nucleic acid may also be used to regulate expression
CC of the PS1/429 gene, and both nucleic acids and peptides are useful as
CC size markers in electrophoresis, chromatography etc. The transgenic
CC animals are used as models for AD, e.g. for testing drugs. Regulators of
CC the PS1/429 gene or polypeptide can be used to treat e.g. AD or diseases
CC involving mitochondrial pathology, apoptosis and neurodegeneration.
CC Typical regulators are antisense sequences, ribozymes, aptamers,
CC synthetic or natural compounds. (II) may also be used to target other
CC coding sequences to particular cellular locations.
SQ Sequence 2764 BP; 715 A; 624 C; 653 G; 772 T;

```
FT                   /*tag= e
FT                   /number= 5
FT   exon            587. .728
FT                   /*tag= f
FT                   /number= 6
FT   exon            729. .796
FT                   /*tag= g
FT                   /number= 7
FT   exon            797. .1017
FT                   /*tag= h
FT                   /number= 8
FT   exon            1018. .1116
FT                   /*tag= i
FT                   /number= 9
FT   exon            1117. .1203
FT                   /*tag= j
FT                   /number= 10
FT   exon            1204. .1377
FT                   /*tag= k
FT                   /number= 11
FT   exon            1378. .1496
FT                   /*tag= l
FT                   /number= 12
FT   exon            1497. .2765
FT                   /*tag= m
FT                   /number= 13
FT   variation       324. .335
FT                   /*tag= n
FT                   /note= "deletion of 12 nucleotides from 3' end of
FT                          exon 4 by alternative splicing, deletes
FT                          Val26-Gln-29 (not critical to PS1
FT                          function)"
FT   variation       1018. .1116
FT                   /*tag= o
FT                   /note= "absence of exon 9 through splicing
FT                          variation, results in Asp-257 changing
FT                          to Ala and fusion of Ala-257 to Thr-291"
FT   mutation        492
FT                   /*tag= p
FT                   /note= "G to C FAD mutation site (Val82Leu)"
FT   mutation        591
FT                   /*tag= q
FT                   /note= "T to C FAD mutation site (Tyr115His)"
FT   mutation        664
FT                   /*tag= r
FT                   /note= "T to C FAD mutation site (Met139Thr)"
FT   mutation        676
FT                   /*tag= s
FT                   /note= "T to C FAD mutation site (Ile143Thr)"
FT   mutation        684
FT                   /*tag= t
FT                   /note= "A to C FAD mutation site (Met146Leu)"
FT   mutation        736
FT                   /*tag= u
FT                   /note= "A to G FAD mutation site (His163Arg)"
FT   mutation        939
FT                   /*tag= v
FT                   /note= "G to A FAD mutation site (Ala231Thr)"
FT   mutation        985
FT                   /*tag= w
FT                   /note= "C to A FAD mutation site (Ala246Glu)"
FT   mutation        1027
FT                   /*tag= x
FT                   /note= "C to T FAD mutation site (Ala260Val)"
FT   mutation        1039
FT                   /*tag= y
FT                   /note= "C to T FAD mutation site (Pro264Leu)"
FT   mutation        1192
FT                   /*tag= z
FT                   /note= "C to T FAD mutation site (Ala285Val)"
FT   mutation        1104
FT                   /*tag= aa
FT                   /note= "C to G FAD mutation site (Leu286Val)"
FT   mutation        1399
FT                   /*tag= ab
FT                   /note= "G to C FAD mutation site (Gly384Ala)"
FT   mutation        1422
FT                   /*tag= ac
FT                   /note= "C to G FAD mutation site (Leu392Val)"
FT   mutation        1477
FT                   /*tag= ad
FT                   /note= "G to A FAD mutation site (Cys410Tyr)"
FT   mutation        1563
FT                   /*tag= ae
FT                   /note= "A to G FAD mutation site (Ile439Val)"
PN   WO9801549-A2.
PD   15-JAN-1998.
PF   04-JUL-1997; CA0475.
PR   02-JAN-1997; US-034590.
PR   05-JUL-1996; US-021673.
PR   12-JUL-1996; US-021700.
PR   08-NOV-1996; US-029895.
PA   (HSCR-) HSC RES & DEV LP.
PA   (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
PI   Fraser PE, Rommens JM, St George-Hyslop PH;
DR   WPI; 98-286355/25.
DR   P-PSDB; W23964.
PT   New isolated mutant presenilin-1 genes - useful for developing
PT   products for use in detection, diagnosis and therapy of Alzheimer's
PT   disease and for drug screening
PS   Disclosure; Page 178-180; 238pp; English.
CC   This cDNA clone, deposited as ATCC 97124, codes for human
CC   presenilin-1 (hPS1, see W23964).  Mutations in the presenilin genes
CC   have been linked to the development in humans of forms of familial
CC   Alzheimer's disease (FAD) and may be causative of other disorders,
CC   e.g. cognitive, intellectual, neurological or physiological
CC   disorders such as cerebral haemorrhage, schizophrenia, depression,
CC   mental retardation and epilepsy.  Isolation of the hPS1 cDNA
CC   followed genetic mapping of the AD3 region at 14q24.3, construction
CC   of a physical contig spanning the AD3 region, transcription mapping
CC   and analysis of candidate genes, and recovery of candidate genes by
CC   RT-PCR from brain mRNA.  In this cDNA, exon 1 is spliced directly
CC   to exon 3.  Other hPS1 sequences (see also V04667) result from
CC   alternative splicing of the mRNA transcript.  A mouse PS1 homologue
CC   (see V04668) and a human presenilin-2 sequence (see V04669) are
CC   also provided, as well as genomic sequences for hPS1 (see
CC   T99661-71).  Use of the nucleic acids and proteins comprising or
CC   derived from the presenilins is made in screening and diagnosing
CC   FAD, identifying and developing therapeutics for treatment of FAD,
CC   and in producing cell lines and transgenic animals useful as models
CC   of FAD.  Nucleic acids (see V04674-80) encoding presenilin
CC   interacting proteins are also provided.
SQ   Sequence   2765 BP;    715 A;    624 C;    652 G;    773 T;


alignment_scores:
            Quality: 1467.00              Length:    466
              Ratio:  3.954                Gaps:      8
 Percent Similarity:  79.614   Percent Identity: 65.665

alignment_block:
 US-08-509-359B-137 x V04666   ..

 Align seg 1/1  to: V04666  from: 1  to: 2765

      24 GluSerProThrProArgSerCysGlnGluGlyArgGlnGlyProGluAs 40
         |||   |||:::|||   |||       ::::::   |||      |||||
     255 GAGTTACCTGCACCGTTGTCCTACTTCCAGAATGCACAGATGTCTGAGGA 304

      40 pGlyGluAsnThrAlaGlnTrpArgSerGlnGluAsnGluGluAspGlyG 57
         |::::::   :::         |||||||||:::::::::::::::   :
     305 CAACCACCTGAGCAATACTGTACGTAGCCAGAATGACAATAGAGAACGGC 354

      57 luGluAspProAspArgTyrValCysSerGlyValPro..... 69
         ::|||      ||||||           |||   |||
     355 AGGAGCACAACGACAGACGG...AGCCTTGGCCACCCTGAGCCATTATCT 401
```

```
CC   Antibodies against the mutant polypeptide can also be used for this
CC   purpose. Vectors containing or expressing a nucleic acid molecule,
CC   protein or antibody specific for mutant PS1 can be administered to a
CC   patient to reduce the likelihood, or delay the onset, of Alzheimer's
CC   disease, e.g. anti-sense RNA expression can be used to decrease
CC   expression of the PS1 peptide. Transgenic animals expressing the
CC   Alzheimer's disease protein can be used to test candidate therapeutics
CC   and to investigate the normal role of PS1. The PS1 peptide may also be
CC   included in pharmaceutical compositions (vaccines) for Alzheimer's
CC   disease therapy.
SQ   Sequence   2765 BP;    715 A;    625 C;    652 G;    772 T;


alignment_scores:
           Quality: 1462.00              Length:    466
             Ratio:  3.941                Gaps:      8
 Percent Similarity:  79.614   Percent Identity: 65.451

alignment_block:
 US-08-509-359B-137 x T85332   ..

 Align seg 1/1  to: T85332  from: 1  to: 2765

      24 GluSerProThrProArgSerCysGlnGluGlyArgGlnGlyProGluAs 40
         |||   |||:::|||   |||      ::::::   |||      |||||
     255 GAGTTACCTGCACCGTTGTCCTACTTCCAGAATGCACAGATGTCTGAGGA 304

      40 pGlyGluAsnThrAlaGlnTrpArgSerGlnGluAsnGluGluAspGlyG 57
         |::::::   :::         |||||||||:::::::::::::::   :
     305 CAACCACCTGAGCAATACTGTACGTAGCCAGAATGACAATAGAGAACGGC 354

      57 luGluAspProAspArgTyrValCysSerGlyValPro..... 69
         ::|||      ||||||            |||   |||
     355 AGGAGCACAACGACAGACGG...AGCCTTGGCCACCCTGAGCCATTATCT 401

      70 ...GlyArgProProGly..... 74
            |||||||||   |||
     402 AATGGACGACCCCAGGGTAACTCCCGGCAGGTGGTGGAGCAAGATGAGGA 451

      75 .LeuGluGluGluLeuThrLeuLysTyrGlyAlaLysHisValIleMetL 91
          :::||||||||||||||||||||||||||||||||||||||||||||||
     452 AGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 501

      91 euPheValProValThrLeuCysMetIleValValValAlaThrIleLys 107
         ||||||||||||||||||||||||:::||||||||||||||||||||||
     502 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAG 551

     108 SerValArgPheTyrThrGluLysAsnGlyGlnLeuIleTyrThrProPh 124
         ||||||   |||||||||:::|||:::||||||||||||||||||||||
     552 TCAGTCAGCTTTTATACCCGGAAGGATGGGCAGCTAATCTATACCCCATT 601

     124 eThrGluAspThrProSerValGlyGlnArgLeuLeuAsnSerValLeuA 141
         ||||||||||||   :::||||||||||||   |||:::|||:::||||
     602 CACAGAAGATACCGAGACTGTGGGCCAGAGAGCCCTGCACTCAATTCTGA 651

     141 snThrLeuIleMetIleSerValIleValValMetThrIlePheLeuVal 157
         ||:::   ||||||||||||||||||||||||||||||:::||||||
     652 ATGCTGCCATCATGATCAGTGTCATTGTTGTCATGACTATCCTCCTGGTG 701

     158 ValLeuTyrLysTyrArgCysTyrLysPheIleHisGlyTrpLeuIleMe 174
         ||||||||||||||||||||||||   ||||||:::|||||||||::
     702 GTTCTGTATAAATACAGGTGCTATAAGGTCATCCATGCCTGGCTTATTAT 751

     174 tSerSerLeuMetLeuLeuPheLeuPheThrTyrIleTyrLeuGlyGluV 191
         :|||||||||:::|||||||||:::|||:::::::|||||||||||||||
     752 ATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGGAAG 801

     191 alLeuLysThrTyrAsnValAlaMetAspTyrProThrLeuLeuLeuThr 207
         ||:::|||||||||||||||||:::|||||   |||:::   |||
     802 TGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTG 851

     208 ValTrpAsnPheGlyAlaValGlyMetValCysIleHisTrpLysGlyPr 224
         :::|||||||||||||:::|||||||||:::   ||||||||||||||||
     852 ATCTGGAATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCC 901

     224 oLeuValLeuGlnGlnAlaTyrLeuIleMetIleSerAlaLeuMetAlaL 241
         ||||   ||||||||||||||||||||||||||||||||||||||||||
     902 ACTTCGACTCCAGCAGGCATATCTCATTATGATTAGTGCCCTCATGGCCC 951

     241 euValPheIleLysTyrLeuProGluTrpSerAlaTrpValIleLeuGly 257
         ||||||||||||||||||||||||||||||:::||||||:::||||||:::
     952 TGGTGTTTATCAAGTACCTCCCTGAATGGACTGCGTGGCTCATCTTGGCT 1001

     258 AlaIleSerValTyrAspLeuValAlaValLeuCysProLysGlyProLe 274
         :::||||||||||||||||||||||||||||||||||||||||||||||||
    1002 GTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTCCGAAAGGTCCACT 1051

     274 uArgMetLeuValGluThrAlaGlnGluArgAsnGluProIlePheProA 291
         ||||||||||||||||||||||||||||||||||   :::||||||||
    1052 TCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTCCAG 1101

     291 laLeuIleTyrSerSerAlaMetValTrpThrValGlyMetAlaLysLeu 307
         ||||||||||||||||:::|||||||||   |||:::|||||||:::
    1102 CTCTCATTTACTCCTCAACAATGGTGTGGTTGGTGAATATGGCAGAAGGA 1151

     308 AspProSerSerGlnGlyAlaLeu......GlnLeuProTyrAspProGl 322
         ||||||:::::|||      :::      :::      ::::::   ||
    1152 GACCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGCATAATGCAGA 1201

     322 u...MetGluGluAspSerTyrAspSerPheGlyGlu.....ProS 335
         |      |||:::::|||   |||:::   :::|||           :
    1202 AAGCACAGAAAGGGAGTCACAAGACACTGTTGCAGAGAATGATGATGGCG 1251

     335 erTyrProGluValPheGluProProLeuThrGlyTyrProGly..... 349
         :::::   |||   :::|||           ::::::   |||
    1252 GGTTCAGTGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCAT 1301

     350 .....GluGluLeu..... 352
                 :::||||||
    1302 CGCTCTACACCTGAGTCACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT 1351

     353 .....GluGluGluGluGluArgGlyValLysLeuGlyLeuGlyA 366
                |||:::   ||||||||||||||||||||||||||||||
    1352 CCTCGCTGGTGAAGACCCAGAGGAAAGGGGAGTAAAACTTGGATTGGGAG 1401

     366 spPheIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySer 382
         ||||||||||||||||||||||||||||||||:::||||||:::|||
    1402 ATTTCATTTTCTACAGTGTTCTGGTTGGTAAAGCCTCAGCAACAGCCAGT 1451

     383 GlyAspTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLe 399
         ||||||||||||||||||:::||||||||||||||||||||||||||||
    1452 GGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCATATTAATTGGTTT 1501

     399 uCysLeuThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAlaL 416
         ||||||||||||||||||||||||:::||||||||||||||||||||||
    1502 GTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAGCTC 1551

     416 euProIleSerIleThrPheGlyLeuIlePheTyrPheSerThrAspAsn 432
         ||||||||||||||||||||||||:::||||||||||:::||||||
    1552 TTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTAT 1601

     433 LeuValArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIle 448
         ||||||:::||||||||||||   ||||||   ||||||:::||||||
    1602 CTTGTACAGCCTTTTATGGACCAATTAGCATTCCATCAATTTTATATC 1649

seq_name: N_Geneseq_36:V29525

seq_documentation_block:
ID   V29525 standard; cDNA; 2764 BP.
AC   V29525;
DT   13-OCT-1998  (first entry)
DE   Homo sapiens PS-1 cDNA.
KW   PS-1; presenilin; presenilin-1; PSP-1; Alzheimer's disease;
KW   serine protease; neurodegeneration; predisposition; diagnosis; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
```

```
seq_documentation_block:
ID   T59535 standard; cDNA; 1762 BP.
AC   T59535;
DT   07-MAY-1997  (first entry)
DE   Human early onset Alzheimer's disease (EOAD) gene.
KW   Early onset Alzheimer's disease; EOAD; neurodegenerative disease;
KW   diagnosis; gene therapy; antisense; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   cds             174. .1577
FT                   /*tag= a
FT   variation       249. .260
FT                   /*tag= b
FT                   /label= VRSQ
FT                   /note= "splice variant"
PN   WO9703086-A1.
PD   30-JAN-1997.
PF   26-JUN-1996; U11064.
PR   13-JUL-1995; US-001142.
PR   18-JUL-1995; US-001501.
PA   (UYSF-) UNIV SOUTH FLORIDA.
PI   Hardy JA;
DR   WPI; 97-118980/11.
DR   P-PSDB; W11839.
PT   Early onset Alzheimer's disease gene - useful for diagnosing a
PT   pre-disposition to Alzheimer's disease
PS   Claim 1; Fig 1; 44pp; English.
CC   A full-length cDNA (T59535) of the early onset Alzheimer's disease
CC   (EOAD) gene sequence codes for a 467-amino acid polypeptide (W11839).
CC   Another full-length cDNA (T59536) of an EOAD splice variant gene
CC   codes for a 463-amino acid polypeptide (W11840).  The 2 sequences
CC   can be used to generate primers and probes for the diagnosis of
CC   predisposition to Alzheimer's disease, esp. EOAD.  They can also be
CC   used for prodn. of EOAD polypeptides in transformed host cells, and
CC   antisense sequences can be used for the treatment of EOAD.
SQ   Sequence   1762 BP;    442 A;    389 C;    430 G;    478 T;


alignment_scores:
           Quality: 1454.00              Length:    466
             Ratio:  3.919                Gaps:      9
 Percent Similarity:  79.614   Percent Identity: 65.451

alignment_block:
 US-08-509-359B-137 x T59535   ..

 Align seg 1/1  to: T59535  from: 1  to: 1762

      24 GluSerProThrProArgSerCysGlnGluGlyArgGlnGlyProGluAs 40
         |||   |||:::|||   |||      ::::::   |||      |||||
     180 GAGTTACCTGCACCGTTGTCCTACTTCCAGAATGCACAGATGTCTGAGGA 229

      40 pGly......GluAsnThrAlaGlnTrpArgSerGlnGluAsnGluGluA 55
         |:::      :::||||||:::          ::::::::::|||:::|||
     230 CAACCACCTGAGCAATACTNNNNNN...NNNNNNAATGACAATAGAGAAC 276

      55 spGlyGluGluAspProAspArgTyrValCysSerGlyValPro...... 69
            |||::::::      |||   :::          |||
     277 GGCAGGAGCACAACGACAGACGGAGCCTTGGCCACCCTGAGCCATTATCT 326

      70 ...GlyArgProProGly.............................. 74
            |||||||||   |||
     327 AATGGACGACCCCAGGGTAACTCCCGGCAGGTGGTGGAGCAAGATGAGGA 376

      75 .LeuGluGluGluLeuThrLeuLysTyrGlyAlaLysHisValIleMetL 91
          :::||||||||||||||||||||||||||||||||||||||||||||||
     377 AGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 426

      91 euPheValProValThrLeuCysMetIleValValValAlaThrIleLys 107
         ||||||||||||||||||||||||:::||||||||||||||||||||||
     427 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAG 476

     108 SerValArgPheTyrThrGluLysAsnGlyGlnLeuIleTyrThrProPh 124
         ||||||   ||||||||||:::|||:::||||||||||||||||||||||
     477 TCAGTCAGCTTTTATACCCGGAAGGATGGGCAGCTAATCTATACCCCATT 526

     124 eThrGluAspThrProSerValGlyGlnArgLeuLeuAsnSerValLeuA 141
         ||||||||||||   :::||||||||||||   |||:::|||:::||||
     527 CACAGAAGATACCGAGACTGTGGGCCAGAGAGCCCTGCACTCAATTCTGA 576

     141 snThrLeuIleMetIleSerValIleValValMetThrIlePheLeuVal 157
         ||:::   ||||||||||||||||||||||||||||||||:::||||||
     577 ATGCTGCCATCATGATCAGTGTCATTGTTGTCATGACTATCCTCCTGGTG 626

     158 ValLeuTyrLysTyrArgCysTyrLysPheIleHisGlyTrpLeuIleMe 174
         ||||||||||||||||||||||||   ||||||:::||||||||::
     627 GTTCTCTATAAATACAGGTGCTATAAGGTCATCCATGCCTGGCTTATTAT 676

     174 tSerSerLeuMetLeuLeuPheLeuPheThrTyrIleTyrLeuGlyGluV 191
         :|||||||||:::||||||||||:::|||:::::|||||||||||||||
     677 ATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGGAAG 726

     191 alLeuLysThrTyrAsnValAlaMetAspTyrProThrLeuLeuLeuThr 207
         ||:::|||||||||||||||||:::||||||   |||:::   |||
     727 TGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTG 776

     208 ValTrpAsnPheGlyAlaValGlyMetValCysIleHisTrpLysGlyPr 224
         :::||||||||||||:::||||||||||:::   ||||||||||||||||
     777 ATCTGGAATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCC 826

     224 oLeuValLeuGlnGlnAlaTyrLeuIleMetIleSerAlaLeuMetAlaL 241
         ||||   ||||||||||||||||||||||||||||||||||||||||||
     827 ACTTCGACTCCAGCAGGCATATCTCATTATGATTAGTGCCCTCATGGCCC 876

     241 euValPheIleLysTyrLeuProGluTrpSerAlaTrpValIleLeuGly 257
         ||||||||||||||||||||||||||||:::||||||:::||||||:::
     877 TGGTGTTTATCAAGTACCTCCCTGAATGGACTGCGTGGCTCATCTTGGCT 926

     258 AlaIleSerValTyrAspLeuValAlaValLeuCysProLysGlyProLe 274
         :::||||||||||||||||||||||||||||||||||||||||||||||||
     927 GTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTCCGAAAGGTCCACT 976

     274 uArgMetLeuValGluThrAlaGlnGluArgAsnGluProIlePheProA 291
         ||||||||||||||||||||||||||||||||||   :::||||||||
     977 TCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTCCAG 1026

     291 laLeuIleTyrSerSerAlaMetValTrpThrValGlyMetAlaLysLeu 307
         ||||||||||||||||:::||||||||   |||:::||||||:::
    1027 CTCTCATTTACTCCTCAACAATGGTGTGGTTGGTGAATATGGCAGAAGGA 1076

     308 AspProSerSerGlnGlyAlaLeu......GlnLeuProTyrAspProGl 322
         ||||||:::::::|||      :::      :::      |||:::   ||
    1077 GACCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAGA 1126

     322 u...MetGluGluAspSerTyrAspSerPheGlyGlu.........ProS 335
         |      |||:::::::|||   |||:::   :::|||            :
    1127 AAGCACAGAAAGGGAGTCACAAGACACTGTTGCAGAGAATGATGATGGCG 1176

     335 erTyrProGluValPheGluProProLeuThrGlyTyrProGly...... 349
         :::::   |||   :::|||              ::::::   |||
    1177 GGTTCAGTGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCAT 1226

     350 ..............................GluGluLeu............ 352
                                        :::|||||||
    1227 CGCTCTACACCTGAGTCACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT 1276

     353 ..........GluGluGluGluGluArgGlyValLysLeuGlyLeuGlyA 366
                 |||:::   ||||||||||||||||||||||||||||||||||
    1277 CCTCGCTGGTGAAGACCCAGAGGAAAGGGGAGTAAAACTTGGATTGGGAG 1326

     366 spPheIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySer 382
         ||||||||||||||||||||||||||||||||:::|||||||:::|||
    1327 ATTTCATTTTCTACAGTGTTCTGGTTGGTAAAGCCTCAGCAACAGCCAGT 1376

     383 GlyAspTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLe 399
```

```
 337 roGluValPheGluProProLeuThrGlyTyrProGly..... 349
       |||   :::|||           ::::::   |||
1171 GTGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCT 1220

 350 ......................GluGluLeu.............. 352
                           :::||||||
1221 ACACCTGAGTCACGAGCTGCTGTCCAGGAACTTTCCAGCAGTATCCTCGC 1270

 353 ....GluGluGluGluGluArgGlyValLysLeuGlyLeuGlyAspPheI 368
         |||:::   ||||||||||||||||||||||||||||||||||||||
1271 TGGTGAAGACCCAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCA 1320

 368 lePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySerGlyAsp 384
     |||||||||||||||||||||||||||||:::||||||:::|||||||||
1321 TTTTCTACAGTGTTCTGGTTGGTAAAGCCTCAGCAACAGCCAGTGGAGAC 1370

 385 TrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLeuCysLe 401
     ||||||||||||:::||||||||||||||||||||||||||||||||||
1371 TGGAACACAACCATAGCCTGTTTCGTAGCCATATTAATTGGTTTGTGCCT 1420

 401 uThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAlaLeuProI 418
     ||||||||||||||||||:::|||||||||||||||||||||||||||||
1421 TACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAGCTCTTCCAA 1470

 418 leSerIleThrPheGlyLeuIlePheTyrPheSerThrAspAsnLeuVal 434
     ||||||||||||||||||:::|||||||||:::||||||   ||||||
1471 TCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTA 1520

 435 ArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIle 448
     :::||||||||||||   ||||||   ||||||:::||||||
1521 CAGCCTTTTATGGACCAATTAGCATTCCATCAATTTTATATC 1562
```

```
OM of: US-08-509-359B-137 to: Issued_Patents_NA:*   out_format : pfs

Date: Mar 19, 2000 11:49 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-1998 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US08509359/runat_17032000_164055_19478/app_query.fasta.1
 -DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=1000000 -USER=US08509359 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT
 -THREADS=1

Search information block:
Query: US-08-509-359B-137
Query length: 448
Database: Issued_Patents_NA:*
Database sequences: 214294
Database length: 59861208
Search time (sec): 26.330000



score_list:
Sequence            Strd Orig    ZScore    EScore Len  ! Documentation  ..
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-136 + 2336.00 4827.08  2.0e-261  2285  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-136 + 2336.00 4827.08  2.0e-261  2285  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-875-972-28 + 2320.50 4794.80  1.2e-259  2276  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-875-972-1 + 1623.50 3347.27  5.2e-179  1417  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-135 + 1468.00 3018.78  1.0e-160  1962  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-135 + 1468.00 3018.78  1.0e-160  1962  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-133 + 1467.00 3012.16  2.4e-160  2791  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-133 + 1467.00 3012.16  2.4e-160  2791  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-1 + 1461.00 2999.65  1.2e-159  2791  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-1 + 1461.00 2999.65  1.2e-159  2791  !
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-670-964-1 + 1454.00 2990.97  3.7e-159  1762  !  S
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-670-964-3 + 1454.00 2991.05  3.6e-159  1750  !  S
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-670-479-24 + 1449.00 2979.47  1.6e-158  1914  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-5 + 1448.00 2971.24  4.6e-158  3087  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-5 + 1448.00 2971.24  4.6e-158  3087  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-875-972-3 + 1438.00 2959.77  2.0e-157  1488  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-3 + 1418.00 2914.72  6.5e-155  1929  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-3 + 1418.00 2914.72  6.5e-155  1929  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-165 + 1166.50 2390.44  1.0e-125  1895  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-165 + 1166.50 2390.44  1.0e-125  1895  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-27 -  400.50  810.32  1.1e-37   489  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-27 -  400.50  810.32  1.1e-37   489  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-6 +  302.00  596.42  8.7e-26   945  !  S
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-160 +  302.00  596.42  8.7e-26   945  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-6 +  302.00  596.42  8.7e-26   945  !  S
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-160 +  302.00  596.42  8.7e-26   945  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-74 +  299.00  598.64  6.6e-26   489  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-74 +  299.00  598.64  6.6e-26   489  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-108 +  250.00  496.16  3.4e-20   500  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-108 +  250.00  496.16  3.4e-20   500  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-63 -  220.00  434.42  9.2e-17   469  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-63 -  220.00  434.42  9.2e-17   469  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-62 -  218.00  429.47  1.7e-16   498  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-62 -  218.00  429.47  1.7e-16   498  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-12 +  210.00  407.77  2.8e-15   736  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-106 +  210.00  414.15  1.2e-15   448  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-155 +  210.00  407.77  2.8e-15   736  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-12 +  210.00  407.77  2.8e-15   736  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-106 +  210.00  414.15  1.2e-15   448  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-155 +  210.00  407.77  2.8e-15   736  !
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-89 -  192.50  376.27  1.6e-13   499  !
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-89 -  192.50  376.27  1.6e-13   499
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-151 +  191.00  364.16  7.6e-13   1003
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-151 +  191.00  364.16  7.6e-13   1003
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-9 +  183.00  340.50  1.6e-11   1726

seq_name: /cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-136

seq_documentation_block:
; Sequence 136, Application US/08967101
; Patent No. 5840540
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/967,101
;      FILING DATE:  10-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  136:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2285 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-967-101-136


alignment_scores:
            Quality: 2336.00              Length:    448
              Ratio:  5.214                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-509-359B-137 x US-08-967-101-136   ..

 Align seg 1/1  to: US-08-967-101-136  from: 1  to: 2285

       1 MetLeuThrPheMetAlaSerAspSerGluGluGluValCysAspGluAr 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     366 ATGCTCACATTCATGGCCTCTGACAGCGAGGAAGAAGTGTGTGATGAGCG 415

      17 gThrSerLeuMetSerAlaGluSerProThrProArgSerCysGlnGluG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     416 GACGTCCCTAATGTCGGCCGAGAGCCCCACGCCGCGCTCCTGCCAGGAGG 465

      34 lyArgGlnGlyProGluAspGlyGluAsnThrAlaGlnTrpArgSerGln 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     466 GCAGGCAGGGCCCAGAGGATGGAGAGAATACTGCCCAGTGGAGAAGCCAG 515
```

```
Align seg 1/1  to: US-08-592-541-136  from: 1  to: 2285

      1 MetLeuThrPheMetAlaSerAspSerGluGluGluValCysAspGluAr 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    366 ATGCTCACATTCATGGCCTCTGACAGCGAGGAAGAAGTGTGTGATGAGCG 415

     17 gThrSerLeuMetSerAlaGluSerProThrProArgSerCysGlnGluG 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    416 GACGTCCCTAATGTCGGCCGAGAGCCCCACGCCGCGCTCCTGCCAGGAGG 465

     34 lyArgGlnGlyProGluAspGlyGluAsnThrAlaGlnTrpArgSerGln 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    466 GCAGGCAGGGCCCAGAGGATGGAGAGAATACTGCCCAGTGGAGAAGCCAG 515

     51 GluAsnGluGluAspGlyGluGluAspProAspArgTyrValCysSerGl 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    516 GAGAACGAGGAGGACGGTGAGGAGGACCCTGACCGCTATGTCTGTAGTGG 565

     67 yValProGlyArgProProGlyLeuGluGluGluLeuThrLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    566 GGTTCCCGGGCGGCCGCCAGGCCTGGAGGAAGAGCTGACCCTCAAATACG 615

     84 lyAlaLysHisValIleMetLeuPheValProValThrLeuCysMetIle 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    616 GAGCGAAGCATGTGATCATGCTGTTTGTGCCTGTCACTCTGTGCATGATC 665

    101 ValValValAlaThrIleLysSerValArgPheTyrThrGluLysAsnGl 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    666 GTGGTGGTAGCCACCATCAAGTCTGTGCGCTTCTACACAGAGAAGAATGG 715

    117 yGlnLeuIleTyrThrProPheThrGluAspThrProSerValGlyGlnA 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    716 ACAGCTCATCTACACGCCATTCACTGAGGACACACCCTCGGTGGGCCAGC 765

    134 rgLeuLeuAsnSerValLeuAsnThrLeuIleMetIleSerValIleVal 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    766 GCCTCCTCAACTCCGTGCTGAACACCCTCATCATGATCAGCGTCATCGTG 815

    151 ValMetThrIlePheLeuValValLeuTyrLysTyrArgCysTyrLysPh 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    816 GTTATGACCATCTTCTTGGTGGTGCTCTACAAGTACCGCTGCTACAAGTT 865

    167 eIleHisGlyTrpLeuIleMetSerSerLeuMetLeuLeuPheLeuPheT 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    866 CATCCATGGCTGGTTGATCATGTCTTCACTGATGCTGCTGTTCCTCTTCA 915

    184 hrTyrIleTyrLeuGlyGluValLeuLysThrTyrAsnValAlaMetAsp 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    916 CCTATATCTACCTTGGGGAAGTGCTCAAGACCTACAATGTGGCCATGGAC 965

    201 TyrProThrLeuLeuLeuThrValTrpAsnPheGlyAlaValGlyMetVa 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    966 TACCCCACCCTCTTGCTGACTGTCTGGAACTTCGGGGCAGTGGGCATGGT 1015

    217 lCysIleHisTrpLysGlyProLeuValLeuGlnGlnAlaTyrLeuIleM 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1016 GTGCATCCACTGGAAGGGCCCTCTGGTGCTGCAGCAGGCCTACCTCATCA 1065

    234 etIleSerAlaLeuMetAlaLeuValPheIleLysTyrLeuProGluTrp 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1066 TGATCAGTGCGCTCATGGCCCTAGTGTTCATCAAGTACCTCCCAGAGTGG 1115

    251 SerAlaTrpValIleLeuGlyAlaIleSerValTyrAspLeuValAlaVa 267
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1116 TCCGCGTGGGTCATCCTGGGCGCCATCTCTGTGTATGATCTCGTGGCTGT 1165

    267 lLeuCysProLysGlyProLeuArgMetLeuValGluThrAlaGlnGluA 284
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1166 GCTGTGTCCCAAAGGGCCTCTGAGAATGCTGGTAGAAACTGCCCAGGAGA 1215

    284 rgAsnGluProIlePheProAlaLeuIleTyrSerSerAlaMetValTrp 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||

   1216 GAAATGAGCCCATATTCCCTGCCCTGATATACTCATCTGCCATGGTGTGG 1265

    301 ThrValGlyMetAlaLysLeuAspProSerSerGlnGlyAlaLeuGlnLe 317
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1266 ACGGTTGGCATGGCGAAGCTGGACCCCTCCTCTCAGGGTGCCCTCCAGCT 1315

    317 uProTyrAspProGluMetGluGluAspSerTyrAspSerPheGlyGluP 334
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1316 CCCCTACGACCCGGAGATGGAAGAAGACTCCTATGACAGTTTTGGGGAGC 1365

    334 roSerTyrProGluValPheGluProProLeuThrGlyTyrProGlyGlu 350
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1366 CTTCATACCCCGAAGTCTTTGAGCCTCCCTTGACTGGCTACCCAGGGGAG 1415

    351 GluLeuGluGluGluGluGluArgGlyValLysLeuGlyLeuGlyAspPh 367
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1416 GAGCTGGAGGAAGAGGAGGAAAGGGGCGTGAAGCTTGGCCTCGGGGACTT 1465

    367 eIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySerGlyA 384
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1466 CATCTTCTACAGTGTGCTGGTGGGCAAGGCGGCTGCCACGGGCAGCGGGG 1515

    384 spTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLeuCys 400
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1516 ACTGGAATACCACGCTGGCCTGCTTCGTGGCCATCCTCATTGGCTTGTGT 1565

    401 LeuThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAlaLeuPr 417
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1566 CTGACCCTCCTGCTGCTTGCTGTGTTCAAGAAGGCGCTGCCCGCCCTCCC 1615

    417 oIleSerIleThrPheGlyLeuIlePheTyrPheSerThrAspAsnLeuV 434
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1616 CATCTCCATCACGTTCGGGCTCATCTTTTACTTCTCCACGGACAACCTGG 1665

    434 alArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIle 448
        |||||||||||||||||||||||||||||||||||||||||||
   1666 TGCGGCCGTTCATGGACACCCTGGCCTCCCATCAGCTCTACATC 1709

seq_name: /cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-875-972-28

seq_documentation_block:
; Sequence 28, Application US/08875972
; Patent No. 5985564
;  GENERAL INFORMATION:
;    APPLICANT:  Huntington Potter and Jinhue Li
;    TITLE OF INVENTION:  ASSAY FOR IDENTIFYING GENES CAUSING
;    TITLE OF INVENTION:  CHROMOSOME NON-DISJUNCTION
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02173-4799
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/875,972
;      FILING DATE:  08-AUG-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/002,448
;      FILING DATE:  16-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan Esq., Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  HU95-03PA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (781) 861-6240
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/875,972
;      FILING DATE:  08-AUG-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/002,448
;      FILING DATE:  16-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan Esq., Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  HU95-03PA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (781) 861-6240
;      TELEFAX:  (781) 861-9540
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1417 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  2..1129
US-08-875-972-1


alignment_scores:
            Quality: 1623.50              Length:    366
              Ratio:  4.861                Gaps:      2
 Percent Similarity:  91.257    Percent Identity: 89.071

alignment_block:
 US-08-509-359B-137 x US-08-875-972-1   ..

 Align seg 1/1  to: US-08-875-972-1  from: 1  to: 1417

      68 ValProGlyArgProProGlyLeuGluGluGluLeuThrLeuLysTyrGl 84
         :::|||   ||||||   ||||||||||||||||||||||||||||||||
          5 ATTCCCTTGCGGCCGACAGGCCTGGAGGAAGAGCTGACCCTCAAATACGG 54

      84 yAlaLysHisValIleMetLeuPheValProValThrLeuCysMetIleV 101
         ||||||||||||||||||||||   ||||||||||||||||||||||||||
         55 AGCGAAGCACGTGATCATGCTGTGTGTGCCTGTCACTCTGTGCATGATCG 104

     101 alValValAlaThrIleLysSerValArgPheTyrThrGluLysAsnGly 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     105 TGGTGGTAGCCACCATCAAGTCTGTGCGCTTCTACACAGAGAAGAATGGA 154

     118 GlnLeuIleTyrThrProPheThrGluAspThrProSerValGlyGlnAr 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     155 CAGCTCATCTACACGCCATTCACTGAGGACACACCCTCGGTGGGCCAGCG 204

     134 gLeuLeuAsnSerValLeuAsnThrLeuIleMetIleSerValIleValV 151
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     205 CCTCCTCAACTCCGTGCTGAACACCCTCATCATGATCAGCGTCATCGTGG 254

     151 alMetThrIlePheLeuValValLeuTyrLysTyrArgCysTyrLysPhe 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     255 TTATGACCATCTTCTTGGTGGTGCTCTACAAGTACCGCTGCTACAAGTTC 304

     168 IleHisGlyTrpLeuIleMetSerSerLeuMetLeuLeuPheLeuPheTh 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     305 ATCCATGGGTGGTTGATCATGTCTTCACTGATGCTGCTGTTCCTCTTCAC 354

     184 rTyrIleTyrLeuGlyGluValLeuLysThrTyrAsnValAlaMetAspT 201
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     355 CTATATCTACCTTGGGGAAGTGCTCAAGACCTACAATGTGGCCATGGACT 404

     201 yrProThrLeuLeuLeuThrValTrpAsnPheGlyAlaValGlyMetVal 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     405 ACCCCACCCTCTTGCTGACTGTCTGGAACTTCGGGGCAGTGGGGCATGGT 454

     218 CysIleHisTrpLysGlyProLeuValLeuGlnGlnAlaTyrLeuIleMe 234
         ||||||||||||||||||||||||||||:::||||||||||||||||||
     455 GTGATCCACTGGAAGGGCCCTCTGGTGCTGGAGCAGGCCTACCTCATCAT 504

     234 tIleSerAlaLeuMetAlaLeuValPheIleLysTyrLeuProGluTrpS 251
         ||||||||||||||||   |||:::||||||||||   ||||||||||
     505 GATCAGTGCGCTCATGCCCCTAATGTTCATCAAGTACCCTCCAGAGTGGT 554

     251 erAlaTrpValIleLeuGlyAlaIleSerValTyrAspLeuValAlaVal 267
         ||||||||||||||||:::   ||||||||||||||||||||:::|||
     555 CCGCGTGGGTGATCCTGGCGCCCATCTCTGTGTATGATCTCGTGACTGTC 604

     268 LeuCysProLysGlyProLeuArgMetLeuValGluThrAlaGlnGluAr 284
         ||||||   ||||||||||||||||||||||||||||||||||||||||
     605 CTGTGTTCCACAGGGCCTCTGAGAATGCTGGTAGAAACTGCCCAGGAGAG 654

     284 gAsnGluProIleProProAlaLeuIleTyrSerSerAlaMetValTrpT 301
         |||||||   ||||||      ||||||||||||||   ||||||||||
     655 AAATGAGACCATATTCTCTCCCCTGATATACTCATCTCCCATGGTGTGGA 704

     301 hrValGlyMetAlaLysLeuAspProSerSerGlnGlyAlaLeuGlnLeu 317
         |||||   |||:::|||||||||||||||||||||||||||||||||||
     705 CGGTTGTCATGTCGAAGCTGGACCCCTCCTCTCAGGGTGCCCTCCAGCTC 754

     318 ProTyrAspProGluMetGluGluAspSerTyrAspSerPheGlyGluPr 334
         ||||||||||||||||||   ||||||||||||||||||||||||||||
     755 CCCTACGACCCGGAGATG...GAAGACTCCTATGACAGTTTTGGGGAGCC 801

     334 oSerTyrProGluValPheGluProProLeuThrGlyTyrProGlyGluG 351
         ||||||||||||||||||||||||||||||:::|||||||||||||||||
     802 TTCATACCCCGAAGTCTTTGAGCCTCCCCTGGCTGGCTACCCAGGGGAGG 851

     351 luLeuGluGluGluGluGluArg......GlyValLysLeuGlyLeuGly 365
         ||||||||||||||||||          ||||||||||||   ||||||
     852 AGCTGGAGGAAGAGGAGGAAAGTCAAGGGGGCGTGAAGCTTGTCCTCGGG 901

     366 AspPheIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySe 382
                            |||      |||:::|||      ||||||||
     902 ACTTCAACTTCCACAGTTGTTCCTGGTGGCCAAGCGCCTCCCACGGGCAG 951

     382 rGlyAspTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyL 399
         ||||||||||   ||||||||||||||||||||||||||||||||||||
     952 CGGGGACTGGATAACCACGCTGGCCTGCTTCGTGGCCATCCTCATTGGCT 1001

     399 euCysLeuThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAla 415
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1002 TGTGTCTGACCCTCCTGCTGCTTGCTGTGTTCAAGAAGGCGCTGCCCGCC 1051

     416 LeuProIleSerIleThrPheGlyLeuIlePheTyrPheSerThrAsp 431
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1052 CTCCCCATCTCCATCACGTTCGGGCTCATCTTTTACTTCTCCACGGAC 1099

seq_name: /cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-967-101-135

seq_documentation_block:
; Sequence 135, Application US/08967101
; Patent No. 5840540
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:. 183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
```

```
    1486 CCCTCCCCATCTCCATCACCTTCGGGCTCGTGTTCTACTTCGCCACGGAT 1535

     432 AsnLeuValArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIl 448
         ||||||:::|||||||||||||   ||||||   ||||||:::|||||
    1536 TACCTTGTGCAGCCCTTCATGGACCAACTTGCATTCCATCAGTTTTATAT 1585

     448 e 448
         |
    1586 C 1586

seq_name: /cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-135

seq_documentation_block:
; Sequence 135, Application US/08592541
; Patent No. 5986054
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  135:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1962 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-592-541-135


alignment_scores:
            Quality: 1468.00              Length:    467
              Ratio:  3.915                Gaps:      6
 Percent Similarity:  80.300   Percent Identity: 64.026

alignment_block:
 US-08-509-359B-137 x US-08-592-541-135   ..

 Align seg 1/1  to: US-08-592-541-135  from: 1  to: 1962

      24 GluSerProThrProArgSerCysGlnGluGlyArgGlnGlyProGluAs 40
         |||   |||:::|||   |||      ::::::   |||      |||||
     192 GAGATACCTGCACCTTTGTCCTACTTCCAGAATGCCCAGATGTCTGAGGA 241

      40 pGlyGluAsnThrAlaGlnTrpArgSerGlnGluAsnGluGluAspGlyG 57
         |:::::::::::::::      |||||||||:::::::::::::::   :
     242 CAGCCACTCCAGCAGCGCCATCCGGAGCCAGAATGACAGCCAAGAACGGC 291

      57 luGluAspProAspArgTyrValCysSerGlyValPro..... 69
         ::::::::   ||||||         ::::::

     292 AGCAGCAGCATGACAGGCAGAGACTTGACAACCCTGAGCCAATATCTAAT 341

      70 GlyArgProProGly.............................Le 75
         |||||||||   :::
     342 GGGCGGCCCCAGAGTAACTCAAGACAGGTGGTGGAACAAGATGAGGAGGA 391

      75 uGluGluGluLeuThrLeuLysTyrGlyAlaLysHisValIleMetLeuP 92
         :::||||||||||||||||||||||||||||||||||||||||||||||
     392 AGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCATGCTCT 441

      92 heValProValThrLeuCysMetIleValValValAlaThrIleLysSer 108
         |||||||||||||||||||||||:::|||||||||||||||||||||||
     442 TTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCA 491

     109 ValArgPheTyrThrGluLysAsnGlyGlnLeuIleTyrThrProPheTh 125
         |||   |||||||||:::|||:::|||||||||||||||||||||||||
     492 GTCAGCTTCTATACCCGGAAGGACGGTCAGCTAATCTACACCCCATTCAC 541

     125 rGluAspThrProSerValGlyGlnArgLeuLeuAsnSerValLeuAsnT 142
         |||||||||   :::||||||||||||   |||:::|||:::|||||||:
     542 AGAAGACACTGAGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATG 591

     142 hrLeuIleMetIleSerValIleValValMetThrIlePheLeuValVal 158
         ::   |||||||||||||||||||||:::|||||||||:::|||||||||
     592 CGGCCATCATGATCAGTGTCATTGTCATTATGACCATCCTCCTGGTGGTC 641

     159 LeuTyrLysTyrArgCysTyrLysPheIleHisGlyTrpLeuIleMetSe 175
         ||||||||||||||||||||||   ||||||:::|||||||||:::||
     642 CTGTATAAATACAGGTGCTACAAGGTCATCCACGCCTGGCTTATTATTTC 691

     175 rSerLeuMetLeuLeuPheLeuPheThrTyrIleTyrLeuGlyGluValL 192
         |||||||:::|||||||||:::|||:::::::|||||||||||||||||:
     692 ATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACTTAGGGGAAGTAT 741

     192 euLysThrTyrAsnValAlaMetAspTyrProThrLeuLeuLeuThrVal 208
         ::|||||||||||||||||:::|||||   |||:::   |||   :::
     742 TTAAGACCTACAATGTCGCCGTGGACTACGTTACAGTAGCACTCCTAATC 791

     209 TrpAsnPheGlyAlaValGlyMetValCysIleHisTrpLysGlyProLe 225
         ||||||||||||:::|||||||||:::::|||||||||||||||||||||
     792 TGGAATTTTGGTGTGGTCGGGATGATTGCCATCCACTGGAAAGGCCCCCT 841

     225 uValLeuGlnGlnAlaTyrLeuIleMetIleSerAlaLeuMetAlaLeuV 242
         |   ||||||||||||||||||||||||||||||||||||||||||||||
     842 TCGACTGCAGCAGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGG 891

     242 alPheIleLysTyrLeuProGluTrpSerAlaTrpValIleLeuGlyAla 258
         |||||||||||||||||||||||||||:::|||||:::|||||||:::::
     892 TATTTATCAAGTACCTCCCCGAATGGACCGCATGGCTCATCTTGGCTGTG 941

     259 IleSerValTyrAspLeuValAlaValLeuCysProLysGlyProLeuAr 275
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     942 ATTTCAGTATATGATTTGGTGGCTGTTTTATGTCCCAAAGGCCCACTTCG 991

     275 gMetLeuValGluThrAlaGlnGluArgAsnGluProIlePheProAlaL 292
         |||||||||||||||||||||||||||||||   :::||||||||||||
    992 TATGCTGGTTGAAACAGCTCAGGAAAGAAATGAGACTCTCTTTCCAGCTC 1041

     292 euIleTyrSerSerAlaMetValTrpThrValGlyMetAlaLysLeuAsp 308
         ||||||||||||:::|||||||||   |||:::||||||:::   |||
    1042 TTATCTATTCCTCAACAATGGTGTGGTTGGTGAATATGGCTGAAGGAGAC 1091

     309 ProSerSerGlnGlyAlaLeuGlnLeuProTyrAspProGluMetGlu.. 324
         |||::::::|||          :::::::|||   :::|||:::   :::
    1092 CCAGAAGCCCAA......AGGAGGGTACCCAAGAACCCCAAGTATAACAC 1135

     325 .....................GluAspSerTyrAspSerPheGlyGluP 334
                               :::|||||||   ::::::      :::
    1136 ACAAAGAGCGGAGAGAGAGACACAGGACAGTGGTTCTGGGAACGATGATG 1185

     334 roSerTyrProGluValPheGluProProLeuThrGlyTyrProGly... 349
         ::::::   |||   :::|||            ::::::   |||
    1186 GTGGCTTCAGTGAGGAGTGGGAGGCCCAAAGAGACAGTCACCTGGGGCCT 1235
```

```
      |||||||||||||||||||||||||||||:::||||||:::||||||:::
    952 TGGTGTTTATCAAGTACCTCCCTGAATGGACTGCGTGGCTCATCTTGGCT 1001

    258 AlaIleSerValTyrAspLeuValAlaValLeuCysProLysGlyProLe 274
        :::||||||||||||||||||||||||||||||||||||||||||||||||
   1002 GTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTCCGAAAGGTCCACT 1051

    274 uArgMetLeuValGluThrAlaGlnGluArgAsnGluProIlePheProA 291
        |||||||||||||||||||||||||||||||||   :::|||||||
   1052 TCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTCCAG 1101

    291 laLeuIleTyrSerSerAlaMetValTrpThrValGlyMetAlaLysLeu 307
        |||||||||||||||||:::||||||||||   |||:::|||||||:::
   1102 CTCTCATTTACTCCTCAACAATGGTGTGGTTGGTGAATATGGCAGAAGGA 1151

    308 AspProSerSerGlnGlyAlaLeu......GlnLeuProTyrAspProGl 322
        ||||||::::::|||      :::      :::      |||:::   ||
   1152 GACCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAGA 1201

    322 u...MetGluGluAspSerTyrAspSerPheGlyGlu.........ProS 335
        |      |||::::::|||   |||:::   :::|||      :
   1202 AAGCACAGAAAGGGAGTCACAAGACACTGTTGCAGAGAATGATGATGGCG 1251

    335 erTyrProGluValPheGluProProLeuThrGlyTyrProGly...... 349
        :::::   |||   :::|||            ::::::   |||
   1252 GGTTCAGTGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCAT 1301

    350 .............................GluGluLeu.......... 352
                                     :::||||||
   1302 CGCTCTACACCTGAGTCACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT 1351

    353 ..........GluGluGluGluGluArgGlyValLysLeuGlyLeuGlyA 366
                ||:::   ||||||||||||||||||||||||||||||
   1352 CCTCGCTGGTGAAGACCCAGAGGAAAGGGGAGTAAAACTTGGATTGGGAG 1401

    366 spPheIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySer 382
        |||||||||||||||||||||||||||||||||:::||||||:::|||
   1402 ATTTCATTTTCTACAGTGTTCTGGTTGGTAAAGCCTCAGCAACAGCCAGT 1451

    383 GlyAspTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLe 399
        |||||||||||||||||:::|||||||||||||||||||||||||||||
   1452 GGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCATATTAATTGGTTT 1501

    399 uCysLeuThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAlaL 416
        ||||||||||||||||||||||:::||||||||||||||||||||||||
   1502 GTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAGCTC 1551

    416 euProIleSerIleThrPheGlyLeuIlePheTyrPheSerThrAspAsn 432
        |||||||||||||||||||||||||:::|||||||||:::||||||
   1552 TTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTAT 1601

    433 LeuValArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIle 448
        ||||||:::|||||||||||||   ||||||   |||||||:::||||||
   1602 CTTGTACAGCCTTTTATGGACCAATTAGCATTCCATCAATTTTATATC 1649

seq_name: /cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-133

seq_documentation_block:
; Sequence 133, Application US/08592541
; Patent No. 5986054
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  133:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2791 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-592-541-133


alignment_scores:
            Quality: 1467.00              Length:    466
              Ratio:  3.954                Gaps:      8
 Percent Similarity:  79.614   Percent Identity: 65.665

alignment_block:
 US-08-509-359B-137 x US-08-592-541-133   ..

 Align seg 1/1  to: US-08-592-541-133  from: 1  to: 2791

     24 GluSerProThrProArgSerCysGlnGluGlyArgGlnGlyProGluAs 40
        |||   |||:::|||   |||      ::::::   |||      |||||
    255 GAGTTACCTGCACCGTTGTCCTACTTCCAGAATGCACAGATGTCTGAGGA 304

     40 pGlyGluAsnThrAlaGlnTrpArgSerGlnGluAsnGluGluAspGlyG 57
        |:::::   :::         |||||||||:::::::::::::::   :
    305 CAACCACCTGAGCAATACTGTACGTAGCCAGAATGACAATAGAGAACGGC 354

     57 luGluAspProAspArgTyrValCysSerGlyValPro..... 69
        ::|||      ||||||            |||   |||
    355 AGGAGCACAACGACAGACGG...AGCCTTGGCCACCCTGAGCCATTATCT 401

     70 ...GlyArgProProGly..................................... 74
           |||||||||   |||
    402 AATGGACGACCCCAGGGTAACTCCCGGCAGGTGGTGGAGCAAGATGAGGA 451

     75 .LeuGluGluGluLeuThrLeuLysTyrGlyAlaLysHisValIleMetL 91
         :::||||||||||||||||||||||||||||||||||||||||||||||
    452 AGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 501

     91 euPheValProValThrLeuCysMetIleValValValAlaThrIleLys 107
        |||||||||||||||||||||||:::|||||||||||||||||||||||
    502 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAG 551

    108 SerValArgPheTyrThrGluLysAsnGlyGlnLeuIleTyrThrProPh 124
        ||||||   |||||||||:::|||:::|||||||||||||||||||||||
    552 TCAGTCAGCTTTTATACCCGGAAGGATGGGCAGCTAATCTATACCCCATT 601

    124 eThrGluAspThrProSerValGlyGlnArgLeuLeuAsnSerValLeuA 141
        ||||||||||||   :::||||||||||||   |||:::|||:::||||
    602 CACAGAAGATACCGAGACTGTGGGCCAGAGAGCCCTGCACTCAATTCTGA 651

    141 snThrLeuIleMetIleSerValIleValValMetThrIlePheLeuVal 157
        ||:::   |||||||||||||||||||||||||||||||:::||||||
    652 ATGCTGCCATCATGATCAGTGTCATTGTTGTCATGACTATCCTCCTGGTG 701

    158 ValLeuTyrLysTyrArgCysTyrLysPheIleHisGlyTrpLeuIleMe 174
        ||||||||||||||||||||||||||   ||||||:::||||||||::
```

```
 75 .LeuGluGluGluLeuThrLeuLysTyrGlyAlaLysHisValIleMetL 91
      :::||||||||||||||||||||||||||||||||||||||||||||||
453 AGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 502

 91 euPheValProValThrLeuCysMetIleValValValAlaThrIleLys 107
    ||||||||||||||||||||||||||:::|||||||||||||||||||||
503 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAG 552

108 SerValArgPheTyrThrGluLysAsnGlyGlnLeuIleTyrThrProPh 124
    ||||||   |||||||||:::|||:::|||||||||||||||||||||||
553 TCAGTCAGCTTTTATACCCGGAAGGATGGGCAGCTAATCTATACCCCATT 602

124 eThrGluAspThrProSerValGlyGlnArgLeuLeuAsnSerValLeuA 141
    ||||||||||||   :::||||||||||||   |||:::|||:::||||
603 CACAGAAGATACCGAGACTGTGGGCCAGAGAGCCCTGCACTCAATTCTGA 652

141 snThrLeuIleMetIleSerValIleValValMetThrIlePheLeuVal 157
    ||:::   |||||||||||||||||||||||||||||||||:::|||||||
653 ATGCTGCCATCATGATCAGTGTCATTGTTGTCATGACTATCCTCCTGGTG 702

158 ValLeuTyrLysTyrArgCysTyrLysPheIleHisGlyTrpLeuIleMe 174
    |||||||||||||||||||||||   ||||||:::||||||||::
703 GTTCTGTATAAATACAGGTGCTATAAGGTCATCCATGCCTGGCTTATTAT 752

174 tSerSerLeuMetLeuLeuPheLeuPheThrTyrIleTyrLeuGlyGluV 191
    :|||||||||:::|||||||||:::|||:::::::|||||||||||||||
753 ATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGGAAG 802

191 alLeuLysThrTyrAsnValAlaMetAspTyrProThrLeuLeuLeuThr 207
    ||:::|||||||||||||||||:::||||||   |||:::   |||
803 TGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTG 852

208 ValTrpAsnPheGlyAlaValGlyMetValCysIleHisTrpLysGlyPr 224
    :::||||||:::|||:::|||||||||:::   ||||||||||||||||
853 ATCTGGAATTTGGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCC 902

224 oLeuValLeuGlnGlnAlaTyrLeuIleMetIleSerAlaLeuMetAlaL 241
    ||||   ||||||||||||||||||||||||||||||||||||||||||||
903 ACTTCGACTCCAGCAGGCATATCTCATTATGATTAGTGCCCTCATGGCCC 952

241 euValPheIleLysTyrLeuProGluTrpSerAlaTrpValIleLeuGly 257
    ||||||||||||||||||||||||||||:::||||||:::|||||||:::
953 TGGTGTTTATCAAGTACCTCCCTGAATGGACTGCGTGGCTCATCTTGGCT 1002

258 AlaIleSerValTyrAspLeuValAlaValLeuCysProLysGlyProLe 274
    :::||||||||||||||||||||||||||||||||||||||||||||||||
1003 GTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTCCGAAAGGTCCACT 1052

274 uArgMetLeuValGluThrAlaGlnGluArgAsnGluProIlePheProA 291
    ||||||||||||||||||||||||||||||||||   :::|||||||
1053 TCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTCCAG 1102

291 laLeuIleTyrSerSerAlaMetValTrpThrValGlyMetAlaLysLeu 307
    ||||||||||||||||:::||||||||||   |||:::|||||||:::
1103 CTCTCATTTACTCCTCAACAATGGTGTGGTTGGTGAATATGGCAGAAGGA 1152

308 AspProSerSerGlnGlyAlaLeu......GlnLeuProTyrAspProGl 322
    ||||||::::::|||      :::      :::      |||:::   ||
1153 GACCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAGA 1202

322 u...MetGluGluAspSerTyrAspSerPheGlyGlu.........ProS 335
    |      |||::::::|||   |||:::   :::|||          :
1203 AAGCACAGAAAGGGAGTCACAAGACACTGTTGCAGAGAATGATGATGGCG 1252

335 erTyrProGluValPheGluProProLeuThrGlyTyrProGly..... 349
    :::::   |||   :::|||            ::::::   |||
1253 GGTTCAGTGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCAT 1302

350 ............................GluGluLeu........... 352
                                :::||||||
1303 CGCTCTACACCTGAGTCACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT 1352

353 .....GluGluGluGluArgGlyValLysLeuGlyLeuGlyA 366
```

```
                |||:::   ||||||||||||||||||||||||||||||
   1353 CCTCGCTGGTGAAGACCCAGAGGAAAGGGGAGTAAAACTTGGATTGGGAG 1402

    366 spPheIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySer 382
        ||||||||||||||||||||||||||||||||||:::|||||||:::|||
   1403 ATTTCATTTTCTACAGTGTTCTGGTTGGTAAAGCCTCAGCAACAGCCAGT 1452

    383 GlyAspTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLe 399
        ||||||||||||||||||:::|||||||||||||||||||||||||||||
   1453 GGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCATATTAATTGGTTT 1502

    399 uCysLeuThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAlaL 416
        ||||||||||||||||||||||||:::||||||||||||||||||||||
   1503 GTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAGCTC 1552

    416 euProIleSerIleThrPheGlyLeuIlePheTyrPheSerThrAspAsn 432
        ||||||||||||||||||||||||:::||||||||||:::||||||
   1553 TTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTAT 1602

    433 LeuValArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIle 448
        ||||||:::||||||||||||   ||||||   ||||||:::||||||
   1603 CTTGTACAGCCTTTTATGGACCAATTAGCATTCCATCAATTTTATATC 1650

seq_name: /cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-1

seq_documentation_block:
; Sequence 1, Application US/08592541
; Patent No. 5986054
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2791 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-592-541-1


alignment_scores:
            Quality: 1461.00              Length:    466
              Ratio:  3.938                Gaps:      8
 Percent Similarity:  79.614    Percent Identity: 65.451

alignment_block:
 US-08-509-359B-137 x US-08-592-541-1   ..
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Han, William T
;      REGISTRATION NUMBER:  34,344
;      REFERENCE/DOCKET NUMBER:  P50358
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5219
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1762 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-670-964-1


alignment_scores:
            Quality: 1454.00              Length:    466
              Ratio:  3.919                Gaps:      9
 Percent Similarity:  79.614   Percent Identity: 65.451

alignment_block:
 US-08-509-359B-137 x US-08-670-964-1   ..

 Align seg 1/1  to: US-08-670-964-1  from: 1  to: 1762

      24 GluSerProThrProArgSerCysGlnGluGlyArgGlnGlyProGluAs 40
         |||   |||:::|||   |||      ::::::   |||      |||||
     180 GAGTTACCTGCACCGTTGTCCTACTTCCAGAATGCACAGATGTCTGAGGA 229

      40 pGly......GluAsnThrAlaGlnTrpArgSerGlnGluAsnGluGluA 55
         |:::      :::|||||:::         :::::::::|||:::|||
     230 CAACCACCTGAGCAATACTNNNNNN...NNNNNNAATGACAATAGAGAAC 276

      55 spGlyGluGluAspProAspArgTyrValCysSerGlyValPro..... 69
            |||::::::      |||   :::             |||
     277 GGCAGGAGCACAACGACAGACGGAGCCTTGGCCACCCTGAGCCATTATCT 326

      70 ...GlyArgProProGly.............................. 74
            |||||||||   |||
     327 AATGGACGACCCCAGGGTAACTCCCGGCAGGTGGTGGAGCAAGATGAGGA 376

      75 .LeuGluGluGluLeuThrLeuLysTyrGlyAlaLysHisValIleMetL 91
          :::||||||||||||||||||||||||||||||||||||||||||||||
     377 AGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 426

      91 euPheValProValThrLeuCysMetIleValValValAlaThrIleLys 107
         |||||||||||||||||||||||||:::|||||||||||||||||||||
     427 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAG 476

     108 SerValArgPheTyrThrGluLysAsnGlyGlnLeuIleTyrThrProPh 124
         ||||||   ||||||||||:::|||:::|||||||||||||||||||||
     477 TCAGTCAGCTTTTATACCCGGAAGGATGGGCAGCTAATCTATACCCCATT 526

     124 eThrGluAspThrProSerValGlyGlnArgLeuLeuAsnSerValLeuA 141
         ||||||||||||   :::||||||||||||   |||:::|||:::||||
     527 CACAGAAGATACCGAGACTGTGGGCCAGAGAGCCCTGCACTCAATTCTGA 576

     141 snThrLeuIleMetIleSerValIleValValMetThrIlePheLeuVal 157
         ||:::   |||||||||||||||||||||||||||||||||:::||||||
     577 ATGCTGCCATCATGATCAGTGTCATTGTTGTCATGACTATCCTCCTGGTG 626

     158 ValLeuTyrLysTyrArgCysTyrLysPheIleHisGlyTrpLeuIleMe 174
         ||||||||||||||||||||||||   ||||||:::||||||||::
     627 GTTCTGTATAAATACAGGTGCTATAAGGTCATCCATGCCTGGCTTATTAT 676

     174 tSerSerLeuMetLeuLeuPheLeuPheThrTyrIleTyrLeuGlyGluV 191
         :|||||||||:::|||||||||:::|||:::::::|||||||||||||||
     677 ATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGGAAG 726

     191 alLeuLysThrTyrAsnValAlaMetAspTyrProThrLeuLeuLeuThr 207
         ||:::|||||||||||||||||||:::||||||   |||:::   |||
     727 TGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTG 776

     208 ValTrpAsnPheGlyAlaValGlyMetValCysIleHisTrpLysGlyPr 224
         :::|||||||||||||:::||||||||||:::   |||||||||||||||||
     777 ATCTGGAATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCC 826

     224 oLeuValLeuGlnGlnAlaTyrLeuIleMetIleSerAlaLeuMetAlaL 241
         ||||   ||||||||||||||||||||||||||||||||||||||||||||
     827 ACTTCGACTCCAGCAGGCATATCTCATTATGATTAGTGCCCTCATGGCCC 876

     241 euValPheIleLysTyrLeuProGluTrpSerAlaTrpValIleLeuGly 257
         ||||||||||||||||||||||||||||||:::||||||:::||||||:::
     877 TGGTGTTTATCAAGTACCTCCCTGAATGGACTGCGTGGCTCATCTTGGCT 926

     258 AlaIleSerValTyrAspLeuValAlaValLeuCysProLysGlyProLe 274
         :::||||||||||||||||||||||||||||||||||||||||||||||||
     927 GTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTCCGAAAGGTCCACT 976

     274 uArgMetLeuValGluThrAlaGlnGluArgAsnGluProIlePheProA 291
         ||||||||||||||||||||||||||||||||||   :::|||||||
     977 TCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTCCAG 1026

     291 laLeuIleTyrSerSerAlaMetValTrpThrValGlyMetAlaLysLeu 307
         ||||||||||||||||:::|||||||||   |||:::|||||||:::
    1027 CTCTCATTTACTCCTCAACAATGGTGTGGTTGGTGAATATGGCAGAAGGA 1076

     308 AspProSerSerGlnGlyAlaLeu......GlnLeuProTyrAspProGl 322
         ||||||::::::|||      :::      :::      |||:::   ||
    1077 GACCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAGA 1126

     322 u...MetGluGluAspSerTyrAspSerPheGlyGlu.........ProS 335
         |      |||::::::|||   |||:::   :::|||           :
    1127 AAGCACAGAAAGGGAGTCACAAGACACTGTTGCAGAGAATGATGATGGCG 1176

     335 erTyrProGluValPheGluProProLeuThrGlyTyrProGly..... 349
         :::::   |||   :::|||            ::::::   |||
    1177 GGTTCAGTGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCAT 1226

     350 .............................GluGluLeu........... 352
                                      :::||||||
    1227 CGCTCTACACCTGAGTCACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT 1276

     353 .........GluGluGluGluGluArgGlyValLysLeuGlyLeuGlyA 366
                 |||:::   ||||||||||||||||||||||||||||||
    1277 CCTCGCTGGTGAAGACCCAGAGGAAAGGGGAGTAAAACTTGGATTGGGAG 1326

     366 spPheIlePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySer 382
         ||||||||||||||||||||||||||||||||||:::||||||:::|||
    1327 ATTTCATTTTCTACAGTGTTCTGGTTGGTAAAGCCTCAGCAACAGCCAGT 1376

     383 GlyAspTrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLe 399
         ||||||||||||||||||:::||||||||||||||||||||||||||||||
    1377 GGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCATATTAATTGGTTT 1426

     399 uCysLeuThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAlaL 416
         ||||||||||||||||||||||||||:::||||||||||||||||||||||
    1427 GTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAGCTC 1476

     416 euProIleSerIleThrPheGlyLeuIlePheTyrPheSerThrAspAsn 432
         ||||||||||||||||||||||||||:::||||||||:::||||||
    1477 TTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTAT 1526

     433 LeuValArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIle 448
         ||||||:::||||||||||||   ||||||   ||||||:::||||||
    1527 CTTGTACAGCCTTTTATGGACCAATTAGCATTCCATCAATTTTATATC 1574

seq_name: /cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-670-964-3

seq_documentation_block:
; Sequence 3, Application US/08670964
; Patent No. 6010874
;  GENERAL INFORMATION:
```

```
      |||||||||||||||||||||||||||||:::||||||:::|||||||||
   1321 TTTTCTACAGTGTTCTGGTTGGTAAAGCCTCAGCAACAGCCAGTGGAGAC 1370

    385 TrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLeuCysLe 401
        ||||||||||||:::||||||||||||||||||||||||||||||||||
   1371 TGGAACACAACCATAGCCTGTTTCGTAGCCATATTAATTGGTTTGTGCCT 1420

    401 uThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAlaLeuProI 418
        ||||||||||||||||||:::|||||||||||||||||||||||||||||
   1421 TACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAGCTCTTCCAA 1470

    418 leSerIleThrPheGlyLeuIlePheTyrPheSerThrAspAsnLeuVal 434
        ||||||||||||||||||||:::|||||||||:::||||||   ||||||
   1471 TCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTA 1520

    435 ArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIle 448
        :::||||||||||||   ||||||   ||||||:::||||||
   1521 CAGCCTTTTATGGACCAATTAGCATTCCATCAATTTTATATC 1562

seq_name: /cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-670-479-24

seq_documentation_block:
; Sequence 24, Application US/08670479
; Patent No. 5973133
;  GENERAL INFORMATION:
;    APPLICANT:  Hardy, John A.
;    APPLICANT:  Goate, Alison M.
;    TITLE OF INVENTION:  MUTANT S182 GENES
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19406-0939
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/670,479
;      FILING DATE:  26-JUN-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/001,500
;      FILING DATE:  18-JUL-1996
;      APPLICATION NUMBER:  60/001,800
;      FILING DATE:  02-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Han, William T
;      REGISTRATION NUMBER:  34,344
;      REFERENCE/DOCKET NUMBER:  P50361
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5219
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1914 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;    FEATURE:
US-08-670-479-24

alignment_scores:
            Quality: 1449.00              Length:    464
              Ratio:  3.916                Gaps:      8
 Percent Similarity:  79.741   Percent Identity: 65.086

alignment_block:
 US-08-509-359B-137 x US-08-670-479-24   ..

 Align seg 1/1  to: US-08-670-479-24  from: 1  to: 1914

      24 GluSerProThrProArgSerCysGlnGluGlyArgGlnGlyProGluAs 40
         |||   |||:::|||   |||      ::::::   |||      |||||
     226 GAGTTACCTGCACCGTTGTCCTACTTCCAGAATGCACAGATGTCTGAGGA 275

      40 pGlyGluAsnThrAlaGlnTrpArgSerGlnGluAsnGluGluAspGlyG 57
         |::::::   :::         ::::::::::::|||:::|||      |
     276 CAACCACCTGAGC.....AATACTAATGACAATAGAGAACGGCAGG 316

      57 luGluAspProAspArgTyrValCysSerGlyValPro.....Gly 70
         ||:::::      |||   :::            |||         |||
     317 AGCACAACGACAGACGGAGCCTTGGCCACCCTGAGCCATTATCTAATGGA 366

      71 ArgProProGly.............................LeuGl 76
         ||||||   |||                              ::
     367 CGACCCCAGGGTAACTCCCGGCAGGTGGTGGAGCAAGATGAGGAAGAAGA 416

      76 uGluGluLeuThrLeuLysTyrGlyAlaLysHisValIleMetLeuPheV 93
         :||||||||||||||||||||||||||||||||||||||||||||||||||
     417 TGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGCTCTTTG 466

      93 alProValThrLeuCysMetIleValValValAlaThrIleLysSerVal 109
         ||||||||||||||||||:::||||||||||||||||||||||||||||
     467 TCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTC 516

     110 ArgPheTyrThrGluLysAsnGlyGlnLeuIleTyrThrProPheThrGl 126
         |||||||||:::|||:::||||||||||||||||||||||||||||||||
     517 AGCTTTTATACCCGGAAGGATGGGCAGCTAATCTATACCCCATTCACAGA 566

     126 uAspThrProSerValGlyGlnArgLeuLeuAsnSerValLeuAsnThrL 143
         |||||||   :::|||||||||||||   |||:::|||:::|||||||:::
     567 AGATACCGAGACTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTG 616

     143 euIleMetIleSerValIleValValMetThrIlePheLeuValValLeu 159
         ||||||||||||||||||||||||||||||||:::|||||||||||||||
     617 CCATCATGATCAGTGTCATTGTTGTCATGACTATCCTCCTGGTGGTTCTG 666

     160 TyrLysTyrArgCysTyrLysPheIleHisGlyTrpLeuIleMetSerSe 176
         ||||||||||||||||||||   ||||||:::||||||||:::||||||
     667 TATAAATACAGGTGCTATAAGGTCATCCATGCCTGGCTTATTATATCATC 716

     176 rLeuMetLeuLeuPheLeuPheThrTyrIleTyrLeuGlyGluValLeuL 193
         ||||:::|||||||||:::|||:::::::|||||||||||||||||:::|
     717 TCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGGAAGTGTTTA 766

     193 ysThrTyrAsnValAlaMetAspTyrProThrLeuLeuLeuThrValTrp 209
         ||||||||||||||||||:::|||||||   |||:::   |||   :::|||
     767 AAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGG 816

     210 AsnPheGlyAlaValGlyMetValCysIleHisTrpLysGlyProLeuVa 226
         |||||||||:::|||||||||:::   ||||||||||||||||||||||
     817 AATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCG 866

     226 lLeuGlnGlnAlaTyrLeuIleMetIleSerAlaLeuMetAlaLeuValP 243
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     867 ACTCCAGCAGGCATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGT 916

     243 heIleLysTyrLeuProGluTrpSerAlaTrpValIleLeuGlyAlaIle 259
         ||||||||||||||||||||||:::||||||:::||||||::::::|||
     917 TTATCAAGTACCTCCCTGAATGGACTGCGTGGCTCATCTTGGCTGTGATT 966

     260 SerValTyrAspLeuValAlaValLeuCysProLysGlyProLeuArgMe 276
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     967 TCAGTATATGATTTAGTGGCTGTTTTGTGTCCGAAAGGTCCACTTCGTAT 1016
```

```
     ||||:::|||||||||:::|||:::::::||||||||||||||||||:::|
1055 TCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGGAAGTGTTTA 1104

 193 ysThrTyrAsnValAlaMetAspTyrProThrLeuLeuLeuThrValTrp 209
     ||||||||||||||||:::|||||||   |||:::   |||   :::|||
1105 AAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGG 1154

 210 AsnPheGlyAlaValGlyMetValCysIleHisTrpLysGlyProLeuVa 226
     |||:::|||:::|||||||||:::   ||||||||||||||||||||
1155 AATTTGGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCG 1204

 226 lLeuGlnGlnAlaTyrLeuIleMetIleSerAlaLeuMetAlaLeuValP 243
      ||||||||||||||||||||||||||||||||||||||||||||||||
1205 ACTCCAGCAGGCATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGT 1254

 243 heIleLysTyrLeuProGluTrpSerAlaTrpValIleLeuGlyAlaIle 259
     ||||||||||||||||||||||:::||||||:::||||||:::::|||
1255 TTATCAAGTACCTCCCTGAATGGACTGCGTGGCTCATCTTGGCTGTGATT 1304

 260 SerValTyrAspLeuValAlaValLeuCysProLysGlyProLeuArgMe 276
     ||||||||||||||||||||||||||||||||||||||||||||||||
1305 TCAGTATATGATTTAGTGGCTGTTTTGTGTCCGAAAGGTCCACTTCGTAT 1354

 276 tLeuValGluThrAlaGlnGluArgAsnGluProIleProProAlaLeuI 293
     ||||||||||||||||||||||||||||||   :::|||||||||||||
1355 GCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTCCAGCTCTCA 1404

 293 leTyrSerSerAlaMetValTrpThrValGlyMetAlaLysLeuAspPro 309
     ||||||||||:::|||||||||   |||:::||||||:::   ||||||
1405 TTTACTCCTCAACAATGGTGTGGTTGGTGAATATGGCAGAAGGAGACCCG 1454

 310 SerSerGlnGlyAlaLeu......GlnLeuProTyrAspProGlu...Me 323
     ::::::|||      :::      :::      |||:::   |||
1455 GAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAGAAAGCAC 1504

 323 tGluGluAspSerTyrAspSerPheGlyGlu.........ProSerTyrP 337
      |||::::::|||   |||:::   :::|||            ::::::
1505 AGAAAGGGAGTCACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCA 1554

 337 roGluValPheGluProProLeuThrGlyTyrProGly............ 349
       |||   :::|||           ::::::   |||
1555 GTGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCT 1604

 350 .......................GluGluLeu................. 352
                            :::||||||
1605 ACACCTGAGTCACGAGCTGCTGTCCAGGAACTTTCCAGCAGTATCCTCGC 1654

 353 ....GluGluGluGluGluArgGlyValLysLeuGlyLeuGlyAspPheI 368
         |||:::   ||||||||||||||||||||||||||||||||||||
1655 TGGTGAAGACCCAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCA 1704

 368 lePheTyrSerValLeuValGlyLysAlaAlaAlaThrGlySerGlyAsp 384
     ||||||||||||||||||||||||||||:::||||||:::|||||||||
1705 TTTTCTACAGTGTTCTGGTTGGTAAAGCCTCAGCAACAGCCAGTGGAGAC 1754

 385 TrpAsnThrThrLeuAlaCysPheValAlaIleLeuIleGlyLeuCysLe 401
     ||||||||||||:::||||||||||||||||||||||||||||||||||
1755 TGGAACACAACCATAGCCTGTTTCGTAGCCATATTAATTGGTTTGTGCCT 1804

 401 uThrLeuLeuLeuLeuAlaValPheLysLysAlaLeuProAlaLeuProI 418
     ||||||||||||||||||:::||||||||||||||||||||||||||||
1805 TACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAGCTCTTCCAA 1854

 418 leSerIleThrPheGlyLeuIlePheTyrPheSerThrAspAsnLeuVal 434
     ||||||||||||||||||:::|||||||||:::||||||   ||||||
1855 TCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTA 1904

 435 ArgProPheMetAspThrLeuAlaSerHisGlnLeuTyrIle 448
     :::||||||||||||   ||||||   ||||||:::||||||
1905 CAGCCTTTTATGGACCAATTAGCATTCCATCAATTTTATATC 1946

seq_name: /cgn2_6/ptodata/2/ina/5D_COMB.seq:US-08-592-541-5
```

```
seq_documentation_block:
; Sequence 5, Application US/08592541
; Patent No. 5986054
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3087 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-592-541-5


alignment_scores:
            Quality: 1448.00              Length:    464
              Ratio:  3.914                Gaps:      8
 Percent Similarity:  79.741   Percent Identity: 65.086

alignment_block:
 US-08-509-359B-137 x US-08-592-541-5   ..

 Align seg 1/1  to: US-08-592-541-5  from: 1  to: 3087

      24 GluSerProThrProArgSerCysGlnGluGlyArgGlnGlyProGluAs 40
         |||   |||:::|||   |||       ::::::   |||      |||||
     564 GAGTTACCTGCACCGTTGTCCTACTTCCAGAATGCACAGATGTCTGAGGA 613

      40 pGlyGluAsnThrAlaGlnTrpArgSerGlnGluAsnGluGluAspGlyG 57
         |::::::   :::         ::::::::::::|||:::|||      |
     614 CAACCACCTGAGC.........AATACTAATGACAATAGAGAACGGCAGG 654

      57 luGluAspProAspArgTyrValCysSerGlyValPro.........Gly 70
         ||::::::      |||   :::          |||         |||
     655 AGCACAACGACAGACGGAGCCTTGGCCACCCTGAGCCATTATCTAATGGA 704

      71 ArgProProGly................................LeuGl 76
         ||||||   |||                                  ::
     705 CGACCCCAGGGTAACTCCCGGCAGGTGGTGGAGCAAGATGAGGAAGAAGA 754

      76 uGluGluLeuThrLeuLysTyrGlyAlaLysHisValIleMetLeuPheV 93
         :|||||||||||||||||||||||||||||||||||||||||||||||||
     755 TGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGCTCTTTG 804

      93 alProValThrLeuCysMetIleValValValAlaThrIleLysSerVal 109
         ||||||||||||||||||:::||||||||||||||||||||||||||||
```

```
OM of: US-08-509-359B-137 to: EST:*   out_format : pfs

Date: Mar 18, 2000 11:52 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-1998 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US08509359/runat_17032000_164054_19458/app_query.fasta.1
 -DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -ALIGN=15 -MODE=LOCAL
 -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08509359
 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-08-509-359B-137
Query length: 448
Database: EST:*
Database sequences: 4538634
Database length: 1887831982
Search time (sec): 350.460000



score_list:
Sequence              Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_est36:AI925372      + 1155.00 2344.25  8.2e-122  703  ! AI925372 wn53d06.x1 NCI_CGAP_Lu
gb_est31:AI675803      + 1120.00 2273.18  7.5e-118  661  ! AI675803 wb97a12.x1 NCI_CGAP_Pr
gb_est35:AI831581      + 1118.00 2268.72  1.3e-117  680  ! AI831581 wj39d04.x1 NCI_CGAP_Lu
gb_est21:AA993681      + 1000.00 2028.34  3.3e-104  587  ! AA993681 ot97b02.s1 Soares_tota
gb_est39:AW131752      +  916.00 1855.74  1.3e-94   597  ! AW131752 xf34b09.x1 NCI_CGAP_Br
gb_est17:AA602396      +  890.00 1803.65  1.1e-91   538  ! AA602396 no30d05.s1 NCI_CGAP_Pr
gb_est35:AI829395      +  877.00 1777.55  3.0e-90   513  ! AI829395 wk64f10.x1 NCI_CGAP_Pa
gb_est33:AI765870      -  870.00 1759.27  3.2e-89   707  ! AI765870 wh66e09.x1 NCI_CGAP_Ki
gb_est44:AW212769      -  847.00 1715.02  9.2e-87   555  ! AW212769 uo66e03.x1 NCI_CGAP_Ma
gb_est8:AA056325       +  835.50 1690.65  2.1e-85   591  ! AA056325 zf53a03.s1 Soares reti
gb_est25:AI276606      +  824.00 1669.57  3.1e-84   481  ! AI276606 ql71h07.x1 Soares_NhHM
gb_est22:AI046485      +  739.00 1494.79  1.7e-74   494  ! AI046485 ud70d05.y1 Sugano mous
gb_est1:T03796         +  732.50 1481.90  8.9e-74   476  ! T03796 IB913 Infant brain, Bent
gb_est23:AI097783      +  704.00 1423.23  1.6e-70   483  ! AI097783 ue35g10.y1 Sugano mous
gb_est36:AI927349      +  704.00 1425.15  1.3e-70   413  ! AI927349 wn49f05.x1 NCI_CGAP_Lu
gb_est24:AI225996      +  671.00 1354.72  1.1e-66   515  ! AI225996 uj08a11.y1 Sugano mous
gb_est11:AA260597      +  663.50 1339.42  7.7e-66   511  ! AA260597 mx76g09.r1 Soares mous
gb_est22:AI039365      +  651.00 1316.98  1.4e-64   393  ! AI039365 ox40a06.s1 Soares_tota
gb_est44:AW177499      +  651.00 1314.92  1.8e-64   465  ! AW177499 PM4-CT0155-270899-001-
gb_est44:AW177510      +  650.00 1312.82  2.3e-64   467  ! AW177510 PM4-CT0155-010999-001-
gb_est20:AA862334      +  621.00 1255.76  3.5e-61   382  ! AA862334 og94a07.s1 NCI_CGAP_Ki
gb_est19:AA782235      +  612.00 1238.29  3.3e-60   352  ! AA782235 ai31c05.s1 Soares_para
gb_est14:AA451661      +  591.00 1193.35  1.1e-57   409  ! AA451661 zx43f07.r1 Soares_tota
gb_est16:AA537185      +  555.00 1114.97  2.4e-53   591  ! AA537185 vk45e03.r1 Soares mous
gb_est11:AA268820      +  543.50 1095.51  3.0e-52   421  ! AA268820 vb01c10.r1 Soares mous
gb_est25:AI324363      +  539.50 1086.90  8.9e-52   435  ! AI324363 mr15d12.y1 Soares mous
gb_est11:AA237206      +  534.00 1074.67  4.3e-51   470  ! AA237206 mx18e01.r1 Soares mous
gb_est10:AA162310      +  495.00  991.87  1.8e-46   589  ! AA162310 mn44a07.r1 Beddington
gb_est5:N24576         +  478.50  961.96  8.1e-45   426  ! N24576 yx72a04.s1 Soares melanc
gb_est11:AA265273      +  475.00  956.11  1.7e-44   382  ! AA265273 mx91c12.r1 Soares mous
gb_est5:H86456         +  474.00  956.11  1.7e-44   323  ! H86456 yt01b06.s1 Soares retina
gb_est26:AA817957      -  474.00  950.50  3.5e-44   511  ! AA817957 UI-R-A0-ag-f-06-0-UI.s
gb_est10:AA144382      +  473.50  950.53  3.5e-44   469  ! AA144382 mr15d12.r1 Soares mous
gb_est25:AI287167      +  472.00  950.48  3.5e-44   366  ! AI287167 ui73h11.y1 Sugano mous
gb_est2:R14600         +  470.50  949.59  4.0e-44   306  ! R14600 yf34b10.r1 Soares fetal
gb_est28:AI528670      -  469.50  943.32  8.9e-44   432  ! AI528670 mr15d12.x1 Soares mous
gb_est5:N27820         +  456.00  913.88  3.9e-42   498  ! N27820 yx54h10.r1 Soares melanc
gb_est2:R05907         +  455.50  918.53  2.1e-42   313  ! R05907 ye93h02.r1 Soares fetal
gb_est11:AA231081      +  423.00  843.04  3.4e-38   642  ! AA231081 mw11d11.r1 Soares mous
gb_est4:H19012         +  421.50  842.57  3.6e-38   519  ! H19012 ym44b02.r1 Soares infant
gb_est11:AA210480      +  417.50  834.50  1.0e-37   513  ! AA210480 mo86b03.r1 Beddington
gb_est18:AA673862      +  413.00  824.06  3.9e-37   566  ! AA673862 vo82h01.r1 Barstead mo
gb_est4:H33787         +  412.00  831.38  1.5e-37   263  ! H33787 EST110123 Rat PC-12 c
gb_est11:AA200706      +  409.00  818.60  7.9e-37   452  ! AA200706 mu03f12.r1 Soares m
gb_est36:AI875963      +  409.00  823.02  4.5e-37   315  ! AI875963 uj53c08.y1 Sugano m

seq_name: gb_est36:AI925372

seq_documentation_block:
LOCUS       AI925372      703 bp    mRNA            EST       02-SEP-1999
DEFINITION  wn53d06.x1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2449163 3'
            similar to SW:PSN2_HUMAN P49810 PRESENILIN 2 ;, mRNA sequence.
ACCESSION   AI925372
VERSION     AI925372.1  GI:5661336
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 703)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On May 18, 1998 this sequence version replaced gi:3137011.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Seq primer: -40UP from Gibco
            High quality sequence stop: 459.
FEATURES             Location/Qualifiers
     source          1. .703
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2449163"
                     /clone_lib="NCI_CGAP_Lu19"
                     /tissue_type="squamous cell carcinoma, poorly
                     differentiated (4 pooled tumors, including primary and
                     metastatic)"
                     /dev_stage="adult"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     pooled lung tumor tissue, and was then primed with a Not I
                     - oligo(dT) primer. Double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I  and Eco RI sites of the modified
                     pT7T3 vector.  Library went througb one round of
                     normalization. Library constructed by Bento Soares and M.
                     Fatima Bonaldo. "
BASE COUNT      138 a    203 c    177 g    184 t      1 others
ORIGIN


alignment_scores:
            Quality: 1155.00             Length:    233
              Ratio:  5.088                Gaps:      0
 Percent Similarity:  97.425   Percent Identity: 97.425

alignment_block:
 US-08-509-359B-137 x AI925372   ..

 Align seg 1/1  to: AI925372  from: 1  to: 703

      79 LeuThrLeuLysTyrGlyAlaLysHisValIleMetLeuPheValProVa 95
         ||||||||||||||||   ||||||||||||||||||||||||||||||||
```

```
     195 rAsnValAlaMetAspTyrProThrLeuLeuLeuThrValTrpAsnPheG 212
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     352 CAATGTGGCCATGGACTACCCCACCCTCTTGCTGACTGTCTGGAACTTCG 401

     212 lyAlaValGlyMetValCysIleHisTrpLysGlyProLeuValLeuGln 228
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     402 GGGCAGTGGGCATGGTGTGCATCCACTGGAAGGGCCCTCTGGTGCTGCAG 451

     229 GlnAlaTyrLeuIleMetIleSerAlaLeuMetAlaLeuValPheIleLy 245
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     452 CAGGCCTACCTCATCATGATCAGTGCGCTCATGGCCCTAGTGTTCATCAA 501

     245 sTyrLeuProGluTrpSerAlaTrpValIleLeuGlyAlaIleSerValT 262
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     502 GTACCTCCCAGAGTGGTCCGCGTGGGTCATCCTGGGCGCCATCTCTGTGT 551

     262 yrAspLeuValAlaValLeuCysProLysGlyProLeuArgMetLeuVal 278
         |||||||||||||||||||||||||   |||||||||||||||||||||
     552 ATGATCTCGTGGCTGTGCTGTGTCCCANAGGGCCTCTGAGAATGCTGGTA 601

     279 GluThrAlaGlnGluArgAsnGluProIlePheProAlaLeuIleTyrSe 295
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     602 GAAACTGCCCAGGAGAGAAATGAGCCCATATTCCCTGCCTTGATATACTC 651

     295 rSerAlaMet 298
         ||||:::|||
     652 ATCTGNCATG 661

seq_name: gb_est35:AI831581

seq_documentation_block:
LOCUS       AI831581      680 bp    mRNA            EST       26-AUG-1999
DEFINITION  wj39d04.x1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2405191 3'
            similar to SW:PSN2_HUMAN P49810 PRESENILIN 2 ;, mRNA sequence.
ACCESSION   AI831581
VERSION     AI831581.1  GI:5452252
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Dec 20, 1995 this sequence version replaced gi:1130797.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Seq primer: -40UP from Gibco
            High quality sequence stop: 456.
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2405191"
                     /clone_lib="NCI_CGAP_Lu19"
                     /tissue_type="squamous cell carcinoma, poorly
                     differentiated (4 pooled tumors, including primary and
                     metastatic)"
                     /dev_stage="adult"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     pooled lung tumor tissue, and was then primed with a Not I
                     - oligo(dT) primer. Double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I  and Eco RI sites of the modified
                     pT7T3 vector.  Library went througb one round of
                     normalization. Library constructed by Bento Soares and M.
                     Fatima Bonaldo. "
BASE COUNT      137 a    191 c    173 g    178 t      1 others
ORIGIN


alignment_scores:
            Quality: 1118.00            Length:    226
              Ratio:  5.059               Gaps:      0
 Percent Similarity:  97.788   Percent Identity: 96.903

alignment_block:
 US-08-509-359B-137 x AI831581   ..

 Align seg 1/1  to: AI831581  from: 1  to: 680

      79 LeuThrLeuLysTyrGlyAlaLysHisValIleMetLeuPheValProVa 95
         ||||||||||||||||   :::|||||||||||||||||||||||||||||
       3 CTGACCCTCAAATACTGATCGAAGCACGTGATCATGCTGTTTGTGCCTGT 52

      95 lThrLeuCysMetIleValValValAlaThrIleLysSerValArgPheT 112
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      53 CACTCTGTGCATGATCGTGGTGGTAGCCACCATCAAGTCTGTGCGCTTCT 102

     112 yrThrGluLysAsnGlyGlnLeuIleTyrThrProPheThrGluAspThr 128
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     103 ACACAGAGAAGAATGGACAGCTCATCTACACGCCATTCACTGAGGACACA 152

     129 ProSerValGlyGlnArgLeuLeuAsnSerValLeuAsnThrLeuIleMe 145
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     153 CCCTCGGTGGGCCAGCGCCTCCTCAACTCCGTGCTGAACACCCTCATCAT 202

     145 tIleSerValIleValValMetThrIlePheLeuValValLeuTyrLysT 162
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     203 GATCAGCGTCATCGTGGTTATGACCATCTTCTTGGTGGTGCTCTACAAGT 252

     162 yrArgCysTyrLysPheIleHisGlyTrpLeuIleMetSerSerLeuMet 178
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     253 ACCGCTGCTACAAGTTCATCCATGGCTGGTTGATCATGTCTTCACTGATG 302

     179 LeuLeuPheLeuPheThrTyrIleTyrLeuGlyGluValLeuLysThrTy 195
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     303 CTGCTGTTCCTCTTCACCTATATCTACCTTGGGGAAGTGCTCAAGACCTA 352

     195 rAsnValAlaMetAspTyrProThrLeuLeuLeuThrValTrpAsnPheG 212
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     353 CAATGTGGCCATGGACTACCCCACCCTCTTGCTGACTGTCTGGAACTTCG 402

     212 lyAlaValGlyMetValCysIleHisTrpLysGlyProLeuValLeuGln 228
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     403 GGGCAGTGGGCATGGTGTGCATCCACTGGAAGGGCCCTCTGGTGCTGCAG 452

     229 GlnAlaTyrLeuIleMetIleSerAlaLeuMetAlaLeuValPheIleLy 245
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     453 CAGGCCTACCTCATCATGATCAGTGCGCTCATGGCCCTAGTGTTCATCAA 502

     245 sTyrLeuProGluTrpSerAlaTrpValIleLeuGlyAlaIleSerValT 262
         ||||||||||||||||||||||||||||||||   |||||||||||||
     503 GTACCTCCCAGAGTGGTCCGCGTGGGTCATACTGNGCGCCATCTCTGTGT 552

     262 yrAspLeuValAlaValLeuCysProLysGlyProLeuArgMetLeuVal 278
         |||||||||||||||||||||||||   |||||||||||||||||||||
     553 ATGATCTCGTGGCTGTGCTGTGTACCATAGGGCCTCTGAGAATGCTGGTA 602

     279 GluThrAlaGlnGluArgAsnGluProIlePheProAlaLeuIleTyrSe 295
         ||||||||||||||||||||:::|||||||||||||||||||||||||||
```

```
          High quality sequence stop: 397.
FEATURES             Location/Qualifiers
     source          1. .597
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2619929"
                     /clone_lib="NCI_CGAP_Brn50"
                     /tissue_type="medulloblastoma"
                     /lab_host="DH10B (phage resistant)"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     medulloblastoma tumor tissue, and was then primed with a
                     Not I - oligo(dT) primer. Double-stranded cDNA was ligated
                     to Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector.  This library is normalized. Library
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      119 a    170 c    147 g    161 t
ORIGIN


alignment_scores:
            Quality:  916.00             Length:    199
              Ratio:  4.796                Gaps:      0
 Percent Similarity:  95.980   Percent Identity: 92.462

alignment_block:
 US-08-509-359B-137 x AW131752   ..

 Align seg 1/1  to: AW131752  from: 1  to: 597

      79 LeuThrLeuLysTyrGlyAlaLysHisValIleMetLeuPheValProVa 95
         ||||||||||||||   :::||||||||||||||||||||||||||||||
       2 CTGACCCTCAAATACTG.TCGAAGCACGTGATCATGCTGTTTGTGCCTGT 50

      95 lThrLeuCysMetIleValValValAlaThrIleLysSerValArgPheT 112
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 CACTCTGTGCATGATCGTGGTGGTAGCCACCATCAAGTCTGTGCGCTTCT 100

     112 yrThrGluLysAsnGlyGlnLeuIleTyrThrProPheThrGluAspThr 128
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 ACACAGAGAAGAATGGACAGCTCATCTACACGCCATTCACTGAGGACACA 150

     129 ProSerValGlyGlnArgLeuLeuAsnSerValLeuAsnThrLeuIleMe 145
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 CCCTCGGTGGGCCAGCGCCTCCTCAACTCCGTGCTGAACACCCTCATCAT 200

     145 tIleSerValIleValValMetThrIlePheLeuValValLeuTyrLysT 162
         ||||||||||||||||   ||||||||||||||||||||||||||||||
     201 GATCAGCGTCATCGTGGGTATGACCATCTTCTTGGTGGTGCTCTACAAGT 250

     162 yrArgCysTyrLysPheIleHisGlyTrpLeuIleMetSerSerLeuMet 178
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 ACCGCTGCTACAAGTTCATCCATGGCTGGTTGATCATGTCTTCACTGATG 300

     179 LeuLeuPheLeuPheThrTyrIleTyrLeuGlyGluValLeuLysThrTy 195
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 CTGCTGTTCCTCTTCACCTATATCTACCTTGGGGAAGTGCTCAAGACCTA 350

     195 rAsnValAlaMetAspTyrProThrLeuLeuLeuThrValTrpAsnPheG 212
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 CAATGTGGCCATGGACTACCCCACCCTCTTGCTGACTGTCTGGAACTTCG 400

     212 lyAlaValGlyMetValCysIleHisTrpLysGlyProLeuValLeuGln 228
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 GGGCAGTGGGCATGGTGTGCATCCACTGGAAGGGCCCTCTGGTGCTGCAG 450

     229 GlnAlaTyrLeuIleMetIleSerAlaLeuMetAlaLeuValPheIleLy 245
         :::|||||||||||||||||||||||||||:::|||:::|||||||||
     451 CACGCCTACCTCATCATGATCAGTGCGCTCATGGTCCTATTGTTCATCAA 500

     245 sTyrLeuProGluTrpSerAlaTrpValIleLeuGlyAlaIleSerValT 262
         |   ||||||:::|||   ||||||||||||||:::|||||||||||||
     501 GATCCTACCAAAGTGGTACGCGTGGGTCATCCTGAGCGCCATATCTGTGT 550

     262 yrAspLeuValAlaValLeuCysProLysGlyProLeuArgMetLeu 277
         ||||||||||||||   ||||||   |||   |||:::   ||||||
     551 ATGATCTCGTCGCTTGGCTGTGTACCAAATGGCCTGTGTGAATGCTG 597

seq_name: gb_est17:AA602396

seq_documentation_block:
LOCUS       AA602396      538 bp    mRNA            EST       08-OCT-1997
DEFINITION  no30d05.s1 NCI_CGAP_Pr22 Homo sapiens cDNA clone IMAGE:1102185 3'
            similar to TR:G1244640 G1244640 PRESENILIN I-374. ;, mRNA sequence.
ACCESSION   AA602396
VERSION     AA602396.1  GI:2436374
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 538)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Sep 19, 1997 this sequence version replaced gi:1517341.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Insert Length: 1338   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 404.
FEATURES             Location/Qualifiers
     source          1. .538
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1102185"
                     /clone_lib="NCI_CGAP_Pr22"
                     /sex="male"
                     /tissue_type="normal prostate"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; 1st strand cDNA was prepared
                     from normal prostate bulk tissue, and was then primed with
                     a Not I - oligo(dT) primer. Double-stranded cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I and cloned into the Not I  and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized, and was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      101 a    155 c    138 g    143 t      1 others
ORIGIN


alignment_scores:
            Quality:  890.00             Length:    179
              Ratio:  5.028                Gaps:      0
 Percent Similarity:  98.883   Percent Identity: 98.324

alignment_block:
 US-08-509-359B-137 x AA602396   ..

 Align seg 1/1  to: AA602396  from: 1  to: 538

      94 ProValThrLeuCysMetIleValValValAlaThrIleLysSerValAr 110
         ||||||||||||||||||||||||||||||||||   ||||||||||||
```

```
DEFINITION  wh66e09.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2385736 3'
            similar to SW:PSN2_HUMAN P49810 PRESENILIN 2 ;, mRNA sequence.
ACCESSION   AI765870
VERSION     AI765870.1  GI:5232379
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 707)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Feb 17, 1998 this sequence version replaced gi:2889754.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Seq primer: -40UP from Gibco
            High quality sequence stop: 466.
FEATURES             Location/Qualifiers
     source          1. .707
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2385736"
                     /clone_lib="NCI_CGAP_Kid11"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo. "
BASE COUNT      177 a    197 c    182 g    150 t      1 others
ORIGIN


alignment_scores:
           Quality:  870.00            Length:    186
             Ratio:  4.807               Gaps:      0
 Percent Similarity:  97.312   Percent Identity: 95.161

alignment_block:
 US-08-509-359B-137 x AI765870/rev   ..

 Align seg 1/1  to reverse of: AI765870  from: 1  to: 707

     263 AspLeuValAlaValLeuCysProLysGlyProLeuArgMetLeuValGl 279
         ||||||   ||||||||||||||||||||   ||||||||||||||||
     707 GATCTGGGGGCTGTGCTGTGTCCCAAAGGGCTTCTGAGAATGCTGGTAGA 658

     279 uThrAlaGlnGluArgAsnGluProIlePheProAlaLeuIleTyrSerS 296
         |||||||::::::||||||||||||||||||||||||||||   :::|
     657 AACTGCCAGGAAGAGAAATGAGCCCATATTCCCTGCCCTGATATA.NCAT 609

     296 erAlaMetValTrpThrValGlyMetAlaLysLeuAspProSerSerGln 312
         ||||||||||||||||||||||||:::|||||||||||||   ||||||
     608 CTGCCATGGTGTGGACGGTTGGCATGGGAAAGCTGGACCCTCC.TCTCAG 560

     313 GlyAlaLeuGlnLeuProTyrAspProGluMetGluGluAspSerTyrAs 329
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     559 GGTGCCCTCCAGCTCCCCTACGACCCGGAGATGGAAGAAGACTCCTATGA 510

     329 pSerPheGlyGluProSerTyrProGluValPheGluProProLeuThrG 346
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     509 CAGTTTTGGGGAGCCTTCATACCCCGAAGTCTTTGAGCCTCCCTTGACTG 460

     346 lyTyrProGlyGluGluLeuGluGluGluGluGluArgGlyValLysLeu 362
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     459 GCTACCCAGGGGAGGAGCTGGAGGAAGAGGAGGAAAGGGGCGTGAAGCTT 410

     363 GlyLeuGlyAspPheIlePheTyrSerValLeuValGlyLysAlaAlaAl 379
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     409 GGCCTCGGGGACTTCATCTTCTACAGTGTGCTGGTGGGCAAGGCGGCTGC 360

     379 aThrGlySerGlyAspTrpAsnThrThrLeuAlaCysPheValAlaIleL 396
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     359 CACGGGCAGCGGGGACTGGAATACCACGCTGGCCTGCTTCGTGGCCATCC 310

     396 euIleGlyLeuCysLeuThrLeuLeuLeuLeuAlaValPheLysLysAla 412
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     309 TCATTGGCTTGTGTCTGACCCTCCTGCTGCTTGCTGTGTTCAAGAAGGCG 260

     413 LeuProAlaLeuProIleSerIleThrPheGlyLeuIlePheTyrPheSe 429
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     259 TGCCCCGCCCTCCCCATTTCCATCACGTTCGGGCTCATCTTTTACTTCTC 210

     429 rThrAspAsnLeuValArgProPheMetAspThrLeuAlaSerHisGlnL 446
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     209 CACGGACAACCTGGTGCGGCCGTTCATGGACACCCTGGCCTCCCATCAGC 160

     446 euTyrIle 448
         ||||||||
     159 TCTACATT 152

seq_name: gb_est44:AW212769

seq_documentation_block:
LOCUS       AW212769      555 bp    mRNA            EST       03-DEC-1999
DEFINITION  uo66e03.x1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:2647516 3'
            similar to gb:L42177 Mus musculus S182 protein mRNA, complete cds
            (MOUSE);, mRNA sequence.
ACCESSION   AW212769
VERSION     AW212769.1  GI:6518888
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 555)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Jul 7, 1999 this sequence version replaced gi:5406916.
            Other_ESTs: uo66e03.y1
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            MGI:1027968
            Seq primer: -40UP from Gibco
            High quality sequence stop: 421.
FEATURES             Location/Qualifiers
     source          1. .555
                     /organism="Mus musculus"
```

```
      |||||||||||||||||||||||||||||||||||||||||||||||||||
    3 CCTGTCACTCTGTGCATGATCGTGGTGGTAGCCACCATCAAGTCTGTGCG 52

  110 gPheTyrThrGluLysAsnGlyGlnLeuIleTyrThrProPheThrGluA 127
      |||||||||||||||||||||||||||||||||||||||||||||||||||
   53 CTTCTACACAGAGAAGAATGGACAGCTCATCTACACGCCATTCACTGAGG 102

  127 spThrProSerValGlyGlnArgLeuLeuAsnSerValLeuAsnThrLeu 143
      |||||||||||||||||||||||||||||||||||||||||||||||||||
  103 ACACACCCTCGGTGGGCCAGCGCCTCCTCAACTCCGTGCTGAACACCCTC 152

  144 IleMetIleSerValIleValValMetThrIlePheLeuValValLeuTy 160
      |||||||||||||||||||||||||||||||||||||||||||||||||||
  153 ATCATGATCAGCGTCATCGTGGTTATGACCATCTTCTTGGTGGTGCTCTA 202

  160 rLysTyrArgCysTyrLysPheIleHisGlyTrpLeuIleMetSerSerL 177
      |||||||||||||||||||||||||||||||||||||||||||||||||||
  203 CAAGTACCGCTGCTACAAGTTCATCCATGGCTGGTTGATCATGTCTTCAC 252

  177 euMetLeuLeuPheLeuPheThrTyrIleTyrLeuGlyGluValLeuLys 193
      |||||||||||||||||||||||||||||||||||||||||||||||||||
  253 TGATGCTGCTGTTCCTCTTCACCTATATCTACCTTGGGGAAGTGCTCAAG 302

  194 ThrTyrAsnValAlaMetAspTyrProThrLeuLeuLeuThrValTrpAs 210
      |||||||||||||||||||||||||||||||||||||||||||||||||||
  303 ACCTACAATGTGGCCATGGACTACCCCACCCTCTTGCTGACTGTCTGGAA 352

  210 nPheGlyAlaValGlyMetValCysIleHisTrpLysGlyProLeuValL 227
      |||||||||||||||||||||||||||||||||||||||||||||||||||
  353 CTTCGGGGCAGTGGGCATGGTGTGCATCCACTGGAAGGGCCCTCTGGTGC 402

  227 euGlnGlnAlaTyrLeuIleMetIleSerAlaLeuMetAlaLeuValPhe 243
      |||||||||||||||||||||||||||||||||||||||||||||||||||
  403 TGCAGCAGGCCTACCTCATCATGATCAGTGCGCTCATGGCCCTAGTGTTC 452

  244 IleLysTyrLeuProGluTrpSerAlaTrpValIleLeuGlyAlaIleSe 260
      |||||||||   |||||||||||||||||||||||||      |||   ::
  453 ATCAAGTACTC.CCAGAGTGGTCCGCGTGGGTCATC...TGGGCGCCANC 498

  260 rValTyrAspLeuValAlaValLeuCysProLysGlyProLeuArgMetL 277
      ::::      :::   :::      |||||||||:::   |||||||||
  499 TCTGTGTATGATCCGTNGCTGTNTGTGTCCAAAGGCT...CTGAGAATGT 545

  277 euValGluThrAlaGlnGluArgAsnGluPro 287
            ::::::|||||||||||||||||||
  546 GG...TAGAANTGCCAGGAGAGAAATGAGCCA 574

seq_name: gb_est25:AI276606

seq_documentation_block:
LOCUS       AI276606      481 bp    mRNA            EST       29-JAN-1999
DEFINITION  ql71h07.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1877821
            3' similar to SW:PSN2_HUMAN P49810 PRESENILIN 2 ;, mRNA sequence.
ACCESSION   AI276606
VERSION     AI276606.1  GI:3898880
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 481)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2285413.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1769   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 448.
FEATURES             Location/Qualifiers
     source          1. .481
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1877821"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT      100 a    138 c    119 g    124 t
ORIGIN


alignment_scores:
            Quality:  824.00             Length:    160
              Ratio:  5.150                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-509-359B-137 x AI276606   ..

 Align seg 1/1  to: AI276606  from: 1  to: 481

   77 GluGluLeuThrLeuLysTyrGlyAlaLysHisValIleMetLeuPheVa 93
      ||||||||||||||||||||||||||||||||||||||||||||||||||
    2 GAAGAGCTGACCCTCAAATACGGAGCGAAGCATGTGATCATGCTGTTTGT 51

   93 lProValThrLeuCysMetIleValValValAlaThrIleLysSerValA 110
      ||||||||||||||||||||||||||||||||||||||||||||||||||
   52 GCCTGTCACTCTGTGCATGATCGTGGTGGTAGCCACCATCAAGTCTGTGC 101

  110 rgPheTyrThrGluLysAsnGlyGlnLeuIleTyrThrProPheThrGlu 126
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  102 GCTTCTACACAGAGAAGAATGGACAGCTCATCTACACGCCATTCACTGAG 151

  127 AspThrProSerValGlyGlnArgLeuLeuAsnSerValLeuAsnThrLe 143
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  152 GACACACCCTCGGTGGGCCAGCGCCTCCTCAACTCCGTGCTGAACACCCT 201

  143 uIleMetIleSerValIleValValMetThrIlePheLeuValValLeuT 160
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  202 CATCATGATCAGCGTCATCGTGGTTATGACCATCTTCTTGGTGGTGCTCT 251

  160 yrLysTyrArgCysTyrLysPheIleHisGlyTrpLeuIleMetSerSer 176
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  252 ACAAGTACCGCTGCTACAAGTTCATCCATGGCTGGTTGATCATGTCTTCA 301

  177 LeuMetLeuLeuPheLeuPheThrTyrIleTyrLeuGlyGluValLeuLy 193
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  302 CTGATGCTGCTGTTCCTCTTCACCTATATCTACCTTGGGGAAGTGCTCAA 351

  193 sThrTyrAsnValAlaMetAspTyrProThrLeuLeuLeuThrValTrpA 210
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  352 GACCTACAATGTGGCCATGGACTACCCCACCCTCTTGCTGACTGTCTGGA 401

  210 snPheGlyAlaValGlyMetValCysIleHisTrpLysGlyProLeuVal 226
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  402 ACTTCGGGGCAGTGGGCATGGTGTGCATCCACTGGAAGGGCCCTCTGGTG 451

  227 LeuGlnGlnAlaTyrLeuIleMetIleSer 236
      ||||||||||||||||||||||||||||||
```

```
                     /db_xref="ATCC (inhost):86429"
                     /db_xref="GDB:D0S9004E"
                     /db_xref="taxon:9606"
                     /clone="IB913"
                     /clone_lib="Infant brain, Bento Soares"
                     /lab_host="E. coli DH5-alpha"
                     /note="Vector: BA, M13-derived; Site_1: HindIII; Site_2:
                     NotI; The infant brain library, constructed by Bento
                     Soares, Columbia University, was oligo-(dT) primed and
                     directionally cloned into an M13-derived plasmid using
                     total brain mRNA from a 72-day old human female afflicted
                     with spinal muscular atrophy. "
BASE COUNT      100 a    133 c    119 g    122 t      2 others
ORIGIN


alignment_scores:
           Quality:  732.50             Length:    154
             Ratio:  4.883                Gaps:      3
 Percent Similarity:  97.403   Percent Identity: 96.753

alignment_block:
 US-08-509-359B-137 x T03796   ..

 Align seg 1/1  to: T03796  from: 1  to: 476

      78 GluLeuThrLeuLysTyrGlyAlaLysHis.ValIleMetLeuPheValP 94
         ||||||||||||||||||||||||       ||||||||||||||||||
      11 GAGCTGACCCTCAAATACGGAGCG...ANACGTNATCATGCTGTTTGTGC 57

      94 roValThrLeuCysMetIleValValValAlaThrIleLysSerValArg 110
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      58 CTGTCACTCTGTGCATGATCGTGGTGGTAGCCACCATCAAGTCTGTGCGC 107

     111 PheTyrThrGluLysAsnGlyGlnLeuIleTyrThrProPheThrGluAs 127
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     108 TTCTACACAGAGAAGAATGGACAGCTCATCTACACGCCATTCACTGAGGA 157

     127 pThrProSerValGlyGlnArgLeuLeuAsnSerValLeuAsnThrLeuI 144
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     158 CACACCCTCGGTGGGCCAGCGCCTCCTCAACTCCGTGCTGAACACCCTCA 207

     144 leMetIleSerValIleValValMetThrIlePheLeuValValLeuTyr 160
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     208 TCATGATCAGCGTCATCGTGGTTATGACCATCTTCTTGGTGGTGCTCTAC 257

     161 LysTyrArgCysTyrLysPheIleHisGlyTrpLeuIleMetSerSerLe 177
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     258 AAGTACCGCTGCTACAAGTTCATCCATGGCTGGTTGATCATGTCTTCACT 307

     177 uMetLeuLeuPheLeuPheThrTyrIleTyrLeuGlyGluValLeuLysT 194
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     308 GATGCTGCTGTTCCTCTTCACCTATATCTACCTTGGGGAAGTGCTCAAGA 357

     194 hrTyrAsnValAlaMetAspTyrProThrLeuLeuLeuThrValTrp.As 210
         |||||||||||||||||||||||||||||||||||||||||||||||| ||
     358 CCTACAATGTGGCCATGGACTACCCCACCCTCTTGCTGACTGTCTGGGAA 407

     210 nPheGlyAlaValGlyMetValCysIleHisTrpLysGlyProLeuValL 227
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     408 CTTCGGGGCAGTGGGCATGGTGTGCATCCACTGGAAGGGGCCTTTGGTGC 457

     227 euGlnGln 229
         |||||:::
     458 TGCAGAGG 465

seq_name: gb_est23:AI097783

seq_documentation_block:
LOCUS       AI097783      483 bp    mRNA            EST       20-AUG-1998
DEFINITION  ue35g10.y1 Sugano mouse liver mlia Mus musculus cDNA clone
            IMAGE:1482402 5' similar to gb:L42177 Mus musculus S182 protein
            mRNA, complete cds (MOUSE);, mRNA sequence.
ACCESSION   AI097783
VERSION     AI097783.1  GI:3447308
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 483)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     On Dec 18, 1997 this sequence version replaced gi:2339814.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:930758
            Seq primer: custom primer used.
FEATURES             Location/Qualifiers
     source          1. .483
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1482402"
                     /clone_lib="Sugano mouse liver mlia"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a  DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert.  Size selection was
                     performed to exclude fragments <1.5kb.  Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science).  Custom primers  for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
BASE COUNT       94 a    152 c    126 g    111 t
ORIGIN


alignment_scores:
           Quality:  704.00             Length:    150
             Ratio:  4.993                Gaps:      0
 Percent Similarity:  94.000   Percent Identity: 90.667

alignment_block:
 US-08-509-359B-137 x AI097783   ..

 Align seg 1/1  to: AI097783  from: 1  to: 483

     286 GluProIlePheProAlaLeuIleTyrSerSerAlaMetValTrpThrVa 302
         |||   :::   |||      :::      :::|||||||||||||||||
      33 GAGGCATTGTGTCCCTTTGCTCTTCCTGTAGCTGCCATGGTGTGGACGGT 82

     302 lGlyMetAlaLysLeuAspProSerSerGlnGlyAlaLeuGlnLeuProT 319
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      83 GGGCATGGCAAAGCTGGACCCCTCCTCTCAAGGAGCGCTGCAGCTCCCCT 132

     319 yrAspProGluMetGluGluAspSerTyrAspSerPheGlyGluProSer 335
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 20, 2000, 04:21:11 ; Search time 35.25 Seconds
(without alignments)
301.032 Million cell updates/sec

Title: US-08-509-359B-137
Perfect score: 2336
Sequence: 1 MLTFMASDSEEEVCDERTSL.....STDNLVRPFMDTLASHQLYI 448

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A_Geneseq_36:*

Word size : 0

Number of hits that pass the threshold : 188963

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2336	100.0	448	1	W05762	Human presenilin-2
2	2336	100.0	448	1	W11321	Human AD4 protein.
3	2336	100.0	448	1	W23967	Human presenilin-2
4	2332	99.8	448	1	W05763	Presenilin-2 M239V
5	2331	99.8	448	1	W05765	Presenilin-2 I420T
6	2327	99.6	448	1	W05764	Presenilin-2 N141I
7	2320.5	99.3	447	1	W28508	Full AD4/AD3LP seq
8	2136	91.4	414	1	W05766	Presenilin-2 delta
9	1623.5	69.5	376	1	W28506	AD4/AD3LP sequence
10	1468	62.8	467	1	W05735	Murine presenilin.
11	1468	62.8	467	1	W23966	Mouse presenilin-1
12	1467	62.8	467	1	W05733	Presenilin-1-1. Ne
13	1467	62.8	467	1	W41430	PS1/467 protein. D
14	1467	62.8	467	1	W23964	Human presenilin-1
15	1464	62.7	467	1	W05755	Presenilin-1-1 L28
16	1464	62.7	467	1	W05758	Presenilin-1-1 L39
17	1464	62.7	467	1	W05737	Presenilin-1-1 V82
18	1464	62.7	467	1	W05746	Presenilin-1-1 I21
19	1463	62.6	467	1	W05754	Presenilin-1-1 A28
20	1463	62.6	467	1	W05736	Presenilin-1-1 A79
21	1463	62.6	467	1	W05747	Presenilin-1-1 I23
22	1463	62.6	467	1	W05749	Presenilin-1-1 A26
23	1463	62.6	467	1	W41431	Mouse PS1/467 prot
24	1462	62.6	467	1	W05738	Presenilin-1-1 V96
25	1462	62.6	467	1	W05739	Presenilin-1-1 Y11
26	1462	62.6	467	1	W05741	Presenilin-1-1 I14
27	1462	62.6	467	1	W05748	Presenilin-1-1 A24
28	1462	62.6	467	1	W27176	Human S182 gene, P
29	1461	62.5	467	1	W05753	Presenilin-1-1 E28
30	1461	62.5	467	1	W05757	Presenilin-1-1 G38
31	1461	62.5	467	1	W05740	Presenilin-1-1 M13
32	1461	62.5	467	1	W05742	Presenilin-1-1 M14
33	1460	62.5	467	1	W05744	Presenilin-1-1 L17
34	1460	62.5	467	1	W56770	Homo sapiens PS-1.
35	1459	62.5	467	1	W05752	Presenilin-1-1 P26
36	1458	62.4	467	1	W05743	Presenilin-1-1 H16
37	1458	62.4	467	1	W05745	Presenilin-1-1 G20
38	1457	62.4	467	1	W05751	Presenilin-1-1 P26
39	1456	62.3	467	1	W05759	Presenilin-1-1 C41
40	1455	62.3	467	1	W05750	Presenilin-1-1 C26
41	1454	62.2	467	1	W11839	Human early onset
42	1454	62.2	463	1	W11840	Early onset Alzhei
43	1448	62.0	463	1	W05734	Presenilin-1-2. Ne
44	1448	62.0	463	1	W23965	Human presenilin-1
45	1446	61.9	463	1	W42375	Human presenilin I

ALIGNMENTS

RESULT 1
W05762
ID W05762 standard; Protein; 448 AA.
AC W05762;
DT 25-JUL-1997 (first entry)
DE Human presenilin-2.
KW Presenilin-2; human; hPS1-1; hPS1-2; PS-2; integral membrane protein; AD;
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW depression; antibody; gene expression modulator; therapy.
OS Homo sapiens.
PN WO9634099-A2.
PD 31-OCT-1996.
PF 29-APR-1996; CA0263.
PR 28-APR-1995; US-431048.
PR 28-JUN-1995; US-496841.
PR 31-JUL-1995; US-509359.
PA (HSCR-) HSC RES & DEV LP.
PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
PI Fraser PE, Rommens JM, St George-Hyslop PH;
DR WPI; 96-497631/49.
DR N-PSDB; T40031.
PT New presenilin genes - useful for diagnosis, therapy and drug
PT screening of familial Alzheimer's disease, cerebral disorders, etc.
PS Claim 4; Page 148-150; 178pp; English.
CC This sequence represents the wild type human presenilin-2.
CC W05733 and W05734 represent the two different forms of wild type human
CC presenilin-1 (PS-1). The form represented by W05734 results from
CC alternate splicing of the genomic DNA sequence. W05735 represents the
CC coding sequence for wild type mouse PS-1. The presenilins are a family of
CC highly conserved integral membrane proteins with a common structural
CC motif, common alternate splicing patterns, and common mutational hot spot
CC regions. Mutations in PS genes are implicated in familial Alzheimer's
CC disease (AD) and possibly other diseases such as cerebral haemorrhage,
CC schizophrenia, depression etc., so detection of mutations in the DNA
CC encoding these sequences can be used for diagnosis of these diseases.
CC These proteins, or vectors that express them or containing antisense
CC sequences, antibodies selective for mutant forms of these proteins (such
CC as W05736) and modulators of PS gene expression are potentially useful
CC for treatment of AD etc. Transgenic animals are useful as models for drug
CC screening. The antibodies can also be used e.g. for affinity purification
CC and in immunoassays.
SQ Sequence 448 AA;

Query Match 100.0%; Score 2336; DB 1; Length 448;
Best Local Similarity 100.0%; Pred. No. 7.3e-237;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60

Qy      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120

Qy     121 YTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 YTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180

Qy     181 FLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
FT                  /label= TM3-4
FT                  /note= "hydrophilic loop"
FT   Domain         201. .218
FT                  /label= TM4
FT                  /note= "transmembrane domain 4"
FT   Domain         219. .224
FT                  /label= TM4-5
FT                  /note= "hydrophilic loop"
FT   Domain         225. .244
FT                  /label= TM5
FT                  /note= "transmembrane domain 5"
FT   Domain         245. .249
FT                  /label= TM5-6
FT                  /note= "hydrophilic loop"
FT   Domain         250. .268
FT                  /label= TM6
FT                  /note= "transmembrane domain 6"
FT   Domain         269. .387
FT                  /label= TM6-7
FT                  /note= "hydrophilic loop"
FT   Domain         388. .409
FT                  /label= TM8
FT                  /note= "transmembrane domain 8"
FT   Misc_difference 141
FT                  /note= "Asn141Ile mutation site (Claim 19)"
FT   Misc_difference 239
FT                  /note= "Met239Val mutation site (Claim 19)"
FT   Misc_difference 420
FT                  /note= "Ile420Thr mutation site"
PN   WO9801549-A2.
PD   15-JAN-1998.
PF   04-JUL-1997; CA0475.
PR   02-JAN-1997; US-034590.
PR   05-JUL-1996; US-021673.
PR   12-JUL-1996; US-021700.
PR   08-NOV-1996; US-029895.
PA   (HSCR-) HSC RES & DEV LP.
PA   (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
PI   Fraser PE, Rommens JM, St George-Hyslop PH;
DR   WPI; 98-286355/25.
DR   N-PSDB; V04669.
PT   New isolated mutant presenilin-1 genes - useful for developing
PT   products for use in detection, diagnosis and therapy of Alzheimer's
PT   disease and for drug screening
PS   Claim 19; Page 203-204; 238pp; English.
CC   This polypeptide comprises human presenilin-2 (hPS2).  Its amino
CC   acid sequence was deduced from an isolated cDNA clone (see V04669).
CC   Human and murine presenilin-1 sequences are also provided (see
CC   W23964-66).  Mutations in the PS-1 and PS-2 genes are linked to
CC   the development in humans of forms of familial Alzheimer's disease
CC   (FAD) and may be causative of other disorders, e.g. cognitive,
CC   intellectual, neurological or physiological disorders such as
CC   cerebral haemorrhage, schizophrenia, depression, mental retardation
CC   and epilepsy.  Use of the nucleic acids and proteins comprising or
CC   derived from the presenilins is made in screening and diagnosing
CC   FAD, identifying and developing therapeutics for treatment of FAD,
CC   and in producing cell lines and transgenic animals useful as models
CC   of FAD.  Methods for identifying substances that bind to, or
CC   modulate the activity of a presenilin protein, and methods for
CC   identifying substances that affect the interaction of a
CC   presenilin-interacting protein with a presenilin protein are also
CC   disclosed.
SQ   Sequence   448 AA;


  Query Match           100.0%;  Score 2336;  DB 1;  Length 448;
  Best Local Similarity  100.0%;  Pred. No. 7.3e-237;
  Matches  448;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60

Qy      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120

Qy     121 YTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 YTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180

Qy     181 FLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 FLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMA 240

Qy     241 LVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 LVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVW 300

Qy     301 TVGMAKLDPSSQGALQLPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGV 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TVGMAKLDPSSQGALQLPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGV 360

Qy     361 KLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISI 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISI 420

Qy     421 TFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           ||||||||||||||||||||||||||||
Db     421 TFGLIFYFSTDNLVRPFMDTLASHQLYI 448


RESULT   4
W05763
ID   W05763 standard; Protein; 448 AA.
AC   W05763;
DT   25-JUL-1997  (first entry)
DE   Presenilin-2 M239V mutation.
KW   Presenilin-2; human; hPS1-1; hPS1-2; PS-2; integral membrane protein; AD;
KW   familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW   depression; antibody; gene expression modulator; therapy; mutein.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   modified_site   239
FT                   /label= M239V
PN   WO9634099-A2.
PD   31-OCT-1996.
PF   29-APR-1996; CA0263.
PR   28-APR-1995; US-431048.
PR   28-JUN-1995; US-496841.
PR   31-JUL-1995; US-509359.
PA   (HSCR-) HSC RES & DEV LP.
PA   (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
PI   Fraser PE,  Rommens JM,  St George-Hyslop PH;
DR   WPI; 96-497631/49.
PT   New presenilin genes - useful for diagnosis, therapy and drug
PT   screening of familial Alzheimer's disease, cerebral disorders, etc.
PS   Claim 4; Page -; 178pp; English.
CC   W05763-W05766 represent mutated versions of the human presenilin-2
CC   protein (see W05762 for wild type sequence). The presenilins are a family
CC   of highly conserved integral membrane proteins with a common structural
CC   motif, common alternate splicing patterns, and common mutational hot spot
CC   regions. Mutations in PS genes are implicated in familial Alzheimer's
CC   disease (AD) and possibly other diseases such as cerebral haemorrhage,
CC   schizophrenia, depression etc., so detection of mutations in the DNA
CC   encoding the wild type sequences can be used for diagnosis of these
CC   diseases. The wild type proteins, or vectors that express them or
CC   containing antisense sequences, antibodies selective for these mutant
CC   forms of the proteins and modulators of PS gene expression are
CC   potentially useful for treatment of AD etc. Transgenic animals are useful
CC   as models for drug screening. The antibodies can also be used e.g. for
CC   affinity purification and in immunoassays.
SQ   Sequence   448 AA;


  Query Match           99.8%;  Score 2332;  DB 1;  Length 448;
```

```
CC   regions. Mutations in PS genes are implicated in familial Alzheimer's
CC   disease (AD) and possibly other diseases such as cerebral haemorrhage,
CC   schizophrenia, depression etc., so detection of mutations in the DNA
CC   encoding the wild type sequences can be used for diagnosis of these
CC   diseases. The wild type proteins, or vectors that express them or
CC   containing antisense sequences, antibodies selective for these mutant
CC   forms of the proteins and modulators of PS gene expression are
CC   potentially useful for treatment of AD etc. Transgenic animals are useful
CC   as models for drug screening. The antibodies can also be used e.g. for
CC   affinity purification and in immunoassays.
SQ   Sequence   448 AA;

  Query Match             99.6%;  Score 2327;  DB 1;  Length 448;
  Best Local Similarity   99.8%;  Pred. No. 6.4e-236;
  Matches  447;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60

Qy      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120

Qy     121 YTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     121 YTPFTEDTPSVGQRLLNSVLITLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180

Qy     181 FLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 FLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMA 240

Qy     241 LVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 LVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVW 300

Qy     301 TVGMAKLDPSSQGALQLPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGV 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TVGMAKLDPSSQGALQLPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGV 360

Qy     361 KLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISI 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISI 420

Qy     421 TFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           ||||||||||||||||||||||||||||
Db     421 TFGLIFYFSTDNLVRPFMDTLASHQLYI 448


RESULT   7
W28508
ID   W28508 standard; Protein; 447 AA.
AC   W28508;
DT   07-DEC-1997  (first entry)
DE   Full AD4/AD3LP sequence.
KW   AD3; AD4/AD3LP; Alzheimer's disease; chromosome; missegregation;
KW   presenilin; inhibitor; AD; trisomy 21; ss.
OS   Homo sapiens.
PN   WO9707213-A2.
PD   27-FEB-1997.
PF   15-AUG-1996; U13314.
PR   16-AUG-1995; US-002448.
PA   (HARD ) HARVARD COLLEGE.
PI   Li J,  Potter H;
DR   WPI; 97-165297/15.
DR   N-PSDB; T87426.
PT   Identifying genes which cause chromosome missegregation - useful for
PT   identifying causes of and treatments for diseases, e.g. Alzheimer's
PT   disease, cancer and ageing
PS   Claim 29; Fig 29; 77pp; English.
CC   Identifying genes which cause improper chromosome segregation,
CC   screening for inhibitors of chromosome missegregation and processes
CC   caused by genes encoding chromosome missegregation promoters
CC   was exemplified using Alzheimer's disease.  The sequences
CC   given in T87401 to T87426 can be used in the above methods.
CC   It is not clear from the figure legend, the figure and the
CC   disclosure of the specification which sequence of Fig 1 and Fig 28
CC   is the AD4/AD3LP or the AD3 sequence.
SQ   Sequence   447 AA;

  Query Match             99.3%;  Score 2320.5;  DB 1;  Length 447;
  Best Local Similarity   99.8%;  Pred. No. 3.1e-235;
  Matches  447;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60

Qy      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120

Qy     121 YTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 YTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180

Qy     181 FLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 FLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMA 240

Qy     241 LVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 LVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVW 300

Qy     301 TVGMAKLDPSSQGALQLPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGV 360
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     301 TVGMAKLDPSSQGALQLPYDPEM-EDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGV 359

Qy     361 KLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISI 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 KLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISI 419

Qy     421 TFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           ||||||||||||||||||||||||||||
Db     420 TFGLIFYFSTDNLVRPFMDTLASHQLYI 447


RESULT   8
W05766
ID   W05766 standard; Protein; 414 AA.
AC   W05766;
DT   25-JUL-1997  (first entry)
DE   Presenilin-2 delta263-296 mutation.
KW   Presenilin-2; human; hPS1-1; hPS1-2; PS-2; integral membrane protein; AD;
KW   familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW   depression; antibody; gene expression modulator; therapy; mutein.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   misc_difference 263. .264
FT                   /note= "site of 34 residue deletion"
PN   WO9634099-A2.
PD   31-OCT-1996.
PF   29-APR-1996; CA0263.
PR   28-APR-1995; US-431048.
PR   28-JUN-1995; US-496841.
PR   31-JUL-1995; US-509359.
PA   (HSCR-) HSC RES & DEV LP.
PA   (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
PI   Fraser PE,  Rommens JM,  St George-Hyslop PH;
DR   WPI; 96-497631/49.
PT   New presenilin genes - useful for diagnosis, therapy and drug
PT   screening of familial Alzheimer's disease, cerebral disorders, etc.
PS   Claim 4; Page -; 178pp; English.
CC   W05763-W05766 represent mutated versions of the human presenilin-2
```

```
CC   coding sequence for wild type human PS-2. The presenilins are a family of
CC   highly conserved integral membrane proteins with a common structural
CC   motif, common alternate splicing patterns, and common mutational hot spot
CC   regions. Mutations in PS genes are implicated in familial Alzheimer's
CC   disease (AD) and possibly other diseases such as cerebral haemorrhage,
CC   schizophrenia, depression etc., so detection of mutations in the DNA
CC   encoding these sequences can be used for diagnosis of these diseases.
CC   These proteins, or vectors that express them or containing antisense
CC   sequences, antibodies selective for mutant forms of these proteins (such
CC   as W05736) and modulators of PS gene expression are potentially useful
CC   for treatment of AD etc. Transgenic animals are useful as models for drug
CC   screening. The antibodies can also be used e.g. for affinity purification
CC   and in immunoassays.
SQ   Sequence   467 AA;


  Query Match           62.8%;  Score 1468;  DB 1;  Length 467;
  Best Local Similarity  64.0%;  Pred. No. 9.9e-146;
  Matches  299;  Conservative   46;  Mismatches   78;  Indels   44;  Gaps    6;

Qy      24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP----GRPPG----- 75
           | | | |  :  |  ||  :::  ||| : :: ::  ||           |||
Db       3 EIPAPLSYFQNAQMSEDSHSSSAIRSQNDSQERQQQHDRQRLDNPEPISNGRPQSNSRQV 62

Qy      75 ------LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDT 128
                 :||||||||||||||||||||||:||||||||| ||| |:||||||||||||
Db      63 VEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDT 122

Qy     129 PSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYL 188
            :|||| |:|:||  |||||||:||| |||||||||| || |||:|||:||| |::|||
Db     123 ETVGQRALHSILNAAIMISVIVIMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYL 182

Qy     189 GEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLP 248
           ||| ||||||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||||
Db     183 GEVFKTYNVAVDYVTVALLIWNFGVVGMIAIHWKGPLRLQQAYLIMISALMALVFIKYLP 242

Qy     249 EWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLD 308
           ||:||:||  |||||||||||||||||||||||||||| :|||||||| ||| | ||: |
Db     243 EWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGD 302

Qy     309 PSSQGALQLPYDPEME--------EDSYDSFGEPSYPEVFEPPLTGYPG----------- 350
           | :|   ::| :|:           :||     :  : | :|     : |
Db     303 PEAQ--RRVPKNPKYNTQRAERETQDSGSGNDDGGFSEEWEAQRDSHLGPHRSTPESRAA 360

Qy     350 -EEL-------EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCL 401
            :||       |: ||||||||||||||||||||:|| |||||||:||||||||||||
Db     361 VQELSGSILTSEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCL 420

Qy     402 TLLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           ||||||:||||||||||||||:|||:|| ||:|||| || || ||
Db     421 TLLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467


RESULT  11
W23966
ID   W23966 standard; Protein; 467 AA.
AC   W23966;
DT   20-JUL-1998  (first entry)
DE   Mouse presenilin-1 homologue.
KW   Presenilin-1; PS1 gene; mouse; familial Alzheimer's disease; FAD;
KW   cerebral haemorrhage; schizophrenia; depression; epilepsy;
KW   mental retardation; diagnosis; therapy; transgenic animal.
OS   Mus musculus.
FH   Key             Location/Qualifiers
FT   Domain          82. .100
FT                   /label= TM1
FT                   /note= "transmembrane domain 1"
FT   Domain          101. .132
FT                   /label= TM1-2
FT                   /note= "hydrophilic loop"
FT   Domain          133. .154
FT                   /label= TM2
FT                   /note= "transmembrane domain 2"
FT   Domain          155. .163
FT                   /label= TM2-3
FT                   /note= "hydrophilic loop"
FT   Domain          164. .183
FT                   /label= TM3
FT                   /note= "transmembrane domain 3"
FT   Domain          184. .194
FT                   /label= TM3-4
FT                   /note= "hydrophilic loop"
FT   Domain          195. .212
FT                   /label= TM4
FT                   /note= "transmembrane domain 4"
FT   Domain          213. .220
FT                   /label= TM4-5
FT                   /note= "hydrophilic loop"
FT   Domain          221. .238
FT                   /label= TM5
FT                   /note= "transmembrane domain 5"
FT   Domain          239. .243
FT                   /label= TM5-6
FT                   /note= "hydrophilic loop"
FT   Domain          244. .262
FT                   /label= TM6
FT                   /note= "transmembrane domain 6"
FT   Domain          263. .407
FT                   /label= TM6-7
FT                   /note= "hydrophilic loop"
FT   Domain          408. .428
FT                   /label= TM8
FT                   /note= "transmembrane domain 8"
FT   Misc_difference 177
FT                   /note= "Phe177Ser mutation site (Claim 1)"
FT   Misc_difference 439
FT                   /note= "Ile439Val mutation site (Claim 1)"
PN   WO9801549-A2.
PD   15-JAN-1998.
PF   04-JUL-1997; CA0475.
PR   02-JAN-1997; US-034590.
PR   05-JUL-1996; US-021673.
PR   12-JUL-1996; US-021700.
PR   08-NOV-1996; US-029895.
PA   (HSCR-) HSC RES & DEV LP.
PA   (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
PI   Fraser PE, Rommens JM, St George-Hyslop PH;
DR   WPI; 98-286355/25.
DR   N-PSDB; V04668.
PT   New isolated mutant presenilin-1 genes - useful for developing
PT   products for use in detection, diagnosis and therapy of Alzheimer's
PT   disease and for drug screening
PS   Disclosure; Page 199-200; 238pp; English.
CC   This polypeptide comprises the murine presenilin-1 (PS1) homologue.
CC   Its amino acid sequence was deduced from an isolated cDNA clone
CC   (see V04668).  Mutations in the human PS1 and PS2 genes (see
CC   V04666-68) have been linked to the development in humans of forms
CC   of familial Alzheimer's disease (FAD).  All amino acids that are
CC   mutated in analysed FAD pedigrees (see W23964) were conserved in
CC   the murine homologue.  Use of the nucleic acids and proteins
CC   comprising or derived from presenilins can be made in screening and
CC   diagnosing FAD, identifying and developing therapeutics for
CC   treatment of FAD, and in producing cell lines and transgenic
CC   animals useful as models of FAD.  Methods for identifying
CC   substances that bind to, or modulate the activity of a presenilin
CC   protein, and methods for identifying substances that affect the
CC   interaction of a presenilin-interacting protein with a presenilin
CC   protein are also disclosed.
SQ   Sequence   467 AA;


  Query Match           62.8%;  Score 1468;  DB 1;  Length 467;
  Best Local Similarity  64.0%;  Pred. No. 9.9e-146;
  Matches  299;  Conservative   46;  Mismatches   78;  Indels   44;  Gaps    6;

Qy      24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP----GRPPG----- 75
```

```
CC   peptides. Antisense nucleic acid may also be used to regulate expression
CC   of the PS1/429 gene, and both nucleic acids and peptides are useful as
CC   size markers in electrophoresis, chromatography etc. The transgenic
CC   animals are used as models for AD, e.g. for testing drugs. Regulators of
CC   the PS1/429 gene or polypeptide can be used to treat e.g. AD or diseases
CC   involving mitochondrial pathology, apoptosis and neurodegeneration.
CC   Typical regulators are antisense sequences, ribozymes, aptamers,
CC   synthetic or natural compounds. (II) may also be used to target other
CC   coding sequences to particular cellular locations.
SQ   Sequence   467 AA;


  Query Match             62.8%;  Score 1467;  DB 1;  Length 467;
  Best Local Similarity   65.7%;  Pred. No. 1.3e-145;
  Matches  306;  Conservative   39;  Mismatches   79;  Indels   42;  Gaps    8;

Qy      24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP-----GRPPG---- 75
           | | | |  :  |  ||   :   ||| :  : :|  ||    | |     ||| |
Db       3 ELPAPLSYFQNAQMSEDNHLSNTVRSQNDNRERQEHNDRR-SLGHPEPLSNGRPQGNSRQ 61

Qy      75 -------LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTED 127
                 :||||||||||||||||||||||:|||||||| ||| |:||||||||||
Db      62 VVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTED 121

Qy     128 TPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIY 187
           | :|||| |:|:||  |||||||||| |||||||||| || |||:|||:||| |::||
Db     122 TETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIY 181

Qy     188 LGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYL 247
           |||| ||||||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||
Db     182 LGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYL 241

Qy     248 PEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKL 307
           |||:||:||  |||||||||||||||||||||||||||| :|||||||| ||| | ||:
Db     242 PEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEG 301

Qy     308 DPSSQGAL--QLPYDPE-MEEDSYDSFGE---PSYPEVFEPPLTGYPG----------- 350
           || :|  :     |: |  | :| |:  |     : | :|     : |
Db     302 DPEAQRRVSKNSKYNAESTERESQDTVAENDDGGFSEEWEAQRDSHLGPHRSTPESRAAV 361

Qy     350 EEL-------EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLT 402
           :||       |: ||||||||||||||||||||:|| |||||||:||||||||||||
Db     362 QELSSSILAGEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLT 421

Qy     403 LLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           |||||:|||||||||||||:|||:|| ||:|||| || || ||
Db     422 LLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467


RESULT  14
W23964
ID   W23964 standard; Protein; 467 AA.
AC   W23964;
DT   20-JUL-1998  (first entry)
DE   Human presenilin-1.
KW   Presenilin-1; PS1 gene; human; familial Alzheimer's disease; FAD;
KW   cerebral haemorrhage; schizophrenia; depression; epilepsy;
KW   mental retardation; diagnosis; therapy; transgenic animal.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          82. .100
FT                   /label= TM1
FT                   /note= "transmembrane domain 1"
FT   Domain          101. .132
FT                   /label= TM1-2
FT                   /note= "hydrophilic loop"
FT   Domain          133. .154
FT                   /label= TM2
FT                   /note= "transmembrane domain 2"
FT   Domain          155. .163
FT                   /label= TM2-3
FT                   /note= "hydrophilic loop"
FT   Domain          164. .183
FT                   /label= TM3
FT                   /note= "transmembrane domain 3"
FT   Domain          184. .194
FT                   /label= TM3-4
FT                   /note= "hydrophilic loop"
FT   Domain          195. .212
FT                   /label= TM4
FT                   /note= "transmembrane domain 4"
FT   Domain          213. .220
FT                   /label= TM4-5
FT                   /note= "hydrophilic loop"
FT   Domain          221. .238
FT                   /label= TM5
FT                   /note= "transmembrane domain 5"
FT   Domain          239. .243
FT                   /label= TM5-6
FT                   /note= "hydrophilic loop"
FT   Domain          244. .262
FT                   /label= TM6
FT                   /note= "transmembrane domain 6"
FT   Domain          263. .407
FT                   /label= TM6-7
FT                   /note= "hydrophilic loop"
FT   Domain          408. .428
FT                   /label= TM8
FT                   /note= "transmembrane domain 8"
FT   Misc_difference 177
FT                   /note= "Phe117Ser mutation site (Claim 1)"
FT   Misc_difference 439
FT                   /note= "Ile439Val mutation site (Claim 1)"
FT   Misc_difference 257
FT                   /note= "Asp257Ala mutation site, associated with
FT                   residue 258-290 deletion (Claim 1)"
FT   Misc_difference 258. .290
FT                   /note= "residue 258-290 deletion mutant, associated
FT                   with Asp257Ala mutation (Claim 1)"
FT   Misc_difference 143
FT                   /note= "Ile143Thr mutation site (Claim 18)"
FT   Misc_difference 146
FT                   /note= "Met146Leu mutation site (Claim 18)"
FT   Misc_difference 171
FT                   /note= "Leu171Pro mutation site (Claim 18)"
FT   Misc_difference 260
FT                   /note= "Ala260Val mutation site (Claim 18)"
FT   Misc_difference 263
FT                   /note= "Cys263Arg mutation site (Claim 18)"
FT   Misc_difference 264
FT                   /note= "Pro264Leu mutation site (Claim 18)"
FT   Misc_difference 267
FT                   /note= "Pro267Ser mutation site (Claim 18)"
FT   Misc_difference 280
FT                   /note= "Glu280Ala mutation site (Claim 18)"
FT   Misc_difference 280
FT                   /note= "Glu280Gly mutation site (Claim 18)"
FT   Misc_difference 285
FT                   /note= "Ala285Val mutation site (Claim 18)"
FT   Misc_difference 286
FT                   /note= "Leu286Val mutation site (Claim 18)"
FT   Misc_difference 322
FT                   /note= "Leu322Val mutation site (Claim 18)"
FT   Misc_difference 392
FT                   /note= "Leu392Val mutation site (Claim 18)"
FT   Misc_difference 410
FT                   /note= "Cys410Tyr mutation site (Claim 18)"
FT   Misc_difference 79
FT                   /note= "Ala79Xaa mutation site"
FT   Misc_difference 82
FT                   /note= "Val82Leu mutation site"
FT   Misc_difference 96
FT                   /note= "Val96Phe mutation site"
FT   Misc_difference 115
FT                   /note= "Tyr115His mutation site"
FT   Misc_difference 139
```

Matches 305; Conservative 40; Mismatches 79; Indels 42; Gaps 8;

```
Qy      24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP-----GRPPG---- 75
           | | | |  :  |  ||   :   ||| :  : :|  ||    | |     ||| |
Db       3 ELPAPLSYFQNAQMSEDNHLSNTVRSQNDNRERQEHNDRR-SLGHPEPLSNGRPQGNSRQ 61

Qy      75 -------LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTED 127
                  :|||||||||||||||||||||:|||||||| ||| |:||||||||||
Db      62 VVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTED 121

Qy     128 TPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIY 187
           | :|||| |:|:||  |||||||||| |||||||||| || |||:|||:||| |::||
Db     122 TETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIY 181

Qy     188 LGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYL 247
           |||| ||||||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||
Db     182 LGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYL 241

Qy     248 PEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKL 307
           |||:||:||  |||||||||||||||||||||||||| :|||:|||| ||| | ||:
Db     242 PEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPAVIYSSTMVWLVNMAEG 301

Qy     308 DPSSQGAL--QLPYDPE-MEEDSYDSFGE---PSYPEVFEPPLTGYPG------------ 350
           || :|  :     |: |  | :| |:  |     : | :|     : |
Db     302 DPEAQRRVSKNSKYNAESTERESQDTVAENDDGGFSEEWEAQRDSHLGPHRSTPESRAAV 361

Qy     350 EEL-------EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLT 402
           :||       |: ||||||||||||||||||||:|| ||||||:|||||||||||||
Db     362 QELSSSILAGEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLT 421

Qy     403 LLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           |||||:|||||||||||||:|||:|| ||:|||| || || ||
Db     422 LLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467
```

Search completed: March 20, 2000, 05:31:20
Job time: 4209 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 18, 2000, 19:01:42 ; Search time 26.47 Seconds
(without alignments)
225.059 Million cell updates/sec

Title: US-08-509-359B-137
Perfect score: 2336
Sequence: 1 MLTFMASDSEEEVCDERTSL.....STDNLVRPFMDTLASHQLYI 448

Scoring table: BLOSUM62

Searched: 133990 seqs, 13297546 residues

Database : Issued_Patents_AA:*

Word size : 0

Number of hits that pass the threshold : 133990
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS9_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2336	100.0	448	2	US-08-967-101-137	Sequence 137, App
2	2336	100.0	448	2	US-08-592-541-137	Sequence 137, App
3	2320.5	99.3	447	2	US-08-875-972-29	Sequence 29, Appl
4	1923	82.3	372	2	US-08-967-101-138	Sequence 138, App
5	1923	82.3	372	2	US-08-592-541-138	Sequence 138, App
6	1623.5	69.5	376	2	US-08-875-972-2	Sequence 2, Appli
7	1467	62.8	467	2	US-08-967-101-134	Sequence 134, App
8	1467	62.8	467	2	US-08-592-541-134	Sequence 134, App
9	1467	62.8	467	3	US-08-670-964-2	Sequence 2, Appli
10	1461	62.5	467	2	US-08-967-101-2	Sequence 2, Appli
11	1461	62.5	467	2	US-08-592-541-2	Sequence 2, Appli
12	1454	62.2	463	3	US-08-670-964-4	Sequence 4, Appli
13	1438	61.6	407	2	US-08-875-972-4	Sequence 4, Appli
14	1437	61.5	463	2	US-08-670-479-18	Sequence 18, Appl
15	1418	60.7	467	2	US-08-967-101-4	Sequence 4, Appli
16	1418	60.7	467	2	US-08-592-541-4	Sequence 4, Appli
17	1150	49.2	541	2	US-08-967-101-166	Sequence 166, App
18	1150	49.2	541	2	US-08-592-541-166	Sequence 166, App
19	110	4.7	1294	2	US-08-819-288-3	Sequence 3, Appli
20	108	4.6	1321	1	US-08-261-822A-3	Sequence 3, Appli
21	108	4.6	1321	4	PCT-US95-07744A-3	Sequence 3, Appli
22	101	4.3	1334	2	US-08-996-545-2	Sequence 2, Appli
23	100	4.3	400	1	US-08-602-010A-8	Sequence 8, Appli
24	100	4.3	400	1	US-08-680-726A-8	Sequence 8, Appli
25	95	4.1	391	1	US-07-816-283-4	Sequence 4, Appli
26	95	4.1	391	1	US-08-417-103-4	Sequence 4, Appli
27	93	4.0	391	1	US-07-816-283-2	Sequence 2, Appli
28	93	4.0	391	1	US-08-417-103-2	Sequence 2, Appli
29	93	4.0	391	1	US-08-417-103-14	Sequence 14, Appl
30	91	3.9	509	2	US-09-031-392-6	Sequence 6, Appli
31	88.5	3.8	452	1	US-08-117-361C-1	Sequence 1, Appli
32	88.5	3.8	3169	2	US-08-477-451-6	Sequence 6, Appli
33	87	3.7	492	2	US-08-355-844-3	Sequence 3, Appli
34	87	3.7	492	4	PCT-US95-16126-3	Sequence 3, Appli
35	86.5	3.7	1480	1	US-07-637-621-2	Sequence 2, Appli
36	86.5	3.7	1480	1	US-08-136-742A-2	Sequence 2, Appli
37	86.5	3.7	1480	1	US-08-135-809A-2	Sequence 2, Appli
38	86.5	3.7	1480	2	US-08-951-912-2	Sequence 2, Appli
39	86.5	3.7	1479	2	US-08-951-912-4	Sequence 4, Appli
40	86.5	3.7	1480	2	US-08-951-912-6	Sequence 6, Appli
41	86.5	3.7	1480	2	US-08-691-605-2	Sequence 2, Appli
42	86.5	3.7	1480	2	US-08-455-552A-14	Sequence 14, Appl
43	86.5	3.7	1480	4	PCT-US93-11667-2	Sequence 2, Appli
44	86.5	3.7	1480	5	5240846-5	Patent No. 5240846
45	86	3.7	413	2	US-08-808-793-25	Sequence 25, Appl

ALIGNMENTS

```
RESULT 1
US-08-967-101-137
; Sequence 137, Application US/08967101
; Patent No. 5840540
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/967,101
;      FILING DATE:  10-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  137:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  448 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-967-101-137


  Query Match           100.0%;  Score 2336;  DB 2;  Length 448;
  Best Local Similarity  100.0%;  Pred. No. 1e-231;
  Matches  448;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60

Qy      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120
```

US-08-875-972-29

Query Match 99.3%; Score 2320.5; DB 2; Length 447;
Best Local Similarity 99.8%; Pred. No. 4e-230;
Matches 447; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60

Qy      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120

Qy     121 YTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 YTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180

Qy     181 FLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 FLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMA 240

Qy     241 LVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 LVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVW 300

Qy     301 TVGMAKLDPSSQGALQLPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGV 360
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db     301 TVGMAKLDPSSQGALQLPYDPEM-EDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGV 359

Qy     361 KLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISI 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 KLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISI 419

Qy     421 TFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           ||||||||||||||||||||||||||||
Db     420 TFGLIFYFSTDNLVRPFMDTLASHQLYI 447
```

RESULT 4
US-08-967-101-138
; Sequence 138, Application US/08967101
; Patent No. 5840540
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,101
; FILING DATE: 10-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 138:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-967-101-138

Query Match 82.3%; Score 1923; DB 2; Length 372;
Best Local Similarity 100.0%; Pred. No. 2e-189;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      77 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 136
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 EELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLL 60

Qy     137 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 196
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 NSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYN 120

Qy     197 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 256
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 VAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVIL 180

Qy     257 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 316
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQ 240

Qy     317 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 376
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 LPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGVKLGLGDFIFYSVLVGK 300

Qy     377 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 436
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRP 360

Qy     437 FMDTLASHQLYI 448
           ||||||||||||
Db     361 FMDTLASHQLYI 372
```

RESULT 5
US-08-592-541-138
; Sequence 138, Application US/08592541
; Patent No. 5986054
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/967,101
;      FILING DATE:  10-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  134:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  467 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-967-101-134


  Query Match             62.8%;  Score 1467;  DB 2;  Length 467;
  Best Local Similarity   65.7%;  Pred. No. 2e-142;
  Matches  306;  Conservative   39;  Mismatches   79;  Indels   42;  Gaps    8;

Qy      24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP-----GRPPG---- 75
           | | | |  :  |  ||   :   ||| :  : :|  ||    | |     ||| |
Db       3 ELPAPLSYFQNAQMSEDNHLSNTVRSQNDNRERQEHNDRR-SLGHPEPLSNGRPQGNSRQ 61

Qy      75 -------LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTED 127
                 :|||||||||||||||||||||||:||||||||| ||| |:||||||||||
Db      62 VVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTED 121

Qy     128 TPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIY 187
           | :|||| |:|:|| ||||||||||| |||||||||| || |||:|||:||| |::||
Db     122 TETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIY 181

Qy     188 LGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYL 247
           |||| ||||||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||
Db     182 LGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYL 241

Qy     248 PEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKL 307
           |||:||:||  |||||||||||||||||||||||||||| :|||||||| ||| | ||:
Db     242 PEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEG 301

Qy     308 DPSSQGAL--QLPYDPE-MEEDSYDSFGE---PSYPEVFEPPLTGYPG------------ 350
           || :|  :     |: |  | :| |:  |     : | :|     : |
Db     302 DPEAQRRVSKNSKYNAESTERESQDTVAENDDGGFSEEWEAQRDSHLGPHRSTPESRAAV 361

Qy     350 EEL-------EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLT 402
           :||       |: |||||||||||||||||||:|| |||||||:|||||||||||||
Db     362 QELSSSILAGEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLT 421

Qy     403 LLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           |||||:|||||||||||||:|||:|| ||:|||| || || ||
Db     422 LLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467


RESULT   8
US-08-592-541-134
; Sequence 134, Application US/08592541
; Patent No. 5986054
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  134:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  467 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-592-541-134


  Query Match             62.8%;  Score 1467;  DB 2;  Length 467;
  Best Local Similarity   65.7%;  Pred. No. 2e-142;
  Matches  306;  Conservative   39;  Mismatches   79;  Indels   42;  Gaps    8;

Qy      24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP-----GRPPG---- 75
           | | | |  :  |  ||   :   ||| :  : :|  ||    | |     ||| |
Db       3 ELPAPLSYFQNAQMSEDNHLSNTVRSQNDNRERQEHNDRR-SLGHPEPLSNGRPQGNSRQ 61

Qy      75 -------LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTED 127
                 :|||||||||||||||||||||||:||||||||| ||| |:||||||||||
Db      62 VVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTED 121

Qy     128 TPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIY 187
           | :|||| |:|:|| ||||||||||| |||||||||| || |||:|||:||| |::||
Db     122 TETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIY 181

Qy     188 LGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYL 247
           |||| ||||||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||
Db     182 LGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYL 241

Qy     248 PEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKL 307
           |||:||:||  |||||||||||||||||||||||||||| :|||||||| ||| | ||:
Db     242 PEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEG 301

Qy     308 DPSSQGAL--QLPYDPE-MEEDSYDSFGE---PSYPEVFEPPLTGYPG------------ 350
           || :|  :     |: |  | :| |:  |     : | :|     : |
Db     302 DPEAQRRVSKNSKYNAESTERESQDTVAENDDGGFSEEWEAQRDSHLGPHRSTPESRAAV 361

Qy     350 EEL-------EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLT 402
           :||       |: |||||||||||||||||||:|| |||||||:|||||||||||||
Db     362 QELSSSILAGEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLT 421

Qy     403 LLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           |||||:|||||||||||||:|||:|| ||:|||| || || ||
Db     422 LLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467


RESULT   9
US-08-670-964-2
; Sequence 2, Application US/08670964
; Patent No. 6010874
;  GENERAL INFORMATION:
;    APPLICANT:  Hardy, John A.
;    TITLE OF INVENTION:  EARLY ONSET ALZHEIMER'S DISEASE
```

```
Qy    308 DPSSQGAL--QLPYDPE-MEEDSYDSFGE---PSYPEVFEPPLTGYPG------------ 350
          || :|  :     |: |  | :| |:  |     : | :|     : |
Db    302 DPEAQRRVSKNSKYNAESTERESQDTVAENDDGGFSEEWEAQRDSHLGPHRSTPESRAAV 361

Qy    350 EEL-------EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLT 402
          :||       |: ||||||||||||||||||||||:|| |||||||:|||||||||||||
Db    362 QELSSSILAGEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLT 421

Qy    403 LLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
          |||||:|||||||||||||||:|||:|| ||:|||| || || ||
Db    422 LLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467


RESULT  11
US-08-592-541-2
; Sequence 2, Application US/08592541
; Patent No. 5986054
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  467 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-592-541-2


  Query Match             62.5%;  Score 1461;  DB 2;  Length 467;
  Best Local Similarity   65.5%;  Pred. No. 8.1e-142;
  Matches  305;  Conservative   39;  Mismatches   80;  Indels   42;  Gaps    8;

Qy     24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP-----GRPPG---- 75
          | | | |  :  |  ||   :   ||| :  : :|  ||    | |     ||| |
Db      3 ELPAPLSYFQNAQMSEDNHLSNTVRSQNDNRERQEHNDRR-SLGHPEPLSNGRPQGNSRQ 61

Qy     75 -------LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTED 127
                 :|||||||||||||||||||||:||||||||| ||| |:|||||||||
Db     62 VVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTED 121

Qy    128 TPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIY 187
          | :|||| |:|:||  |||||||||| |||||||||| || |||:|||:||| |::||
Db    122 TETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIY 181

Qy    188 LGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYL 247
          |||| |||||:|| |: | :|| | |||: ||||||| ||||||||||||||||||||
Db    182 LGEVFKTYNVAVDYITVALLIWNLGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYL 241

Qy    248 PEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKL 307
          |||:||:||  ||||||||||||||||||||||||||| :|||||||| ||| | ||:
Db    242 PEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEG 301

Qy    308 DPSSQGAL--QLPYDPE-MEEDSYDSFGE---PSYPEVFEPPLTGYPG------------ 350
          || :|  :     |: |  | :| |:  |     : | :|     : |
Db    302 DPEAQRRVSKNSKYNAESTERESQDTVAENDDGGFSEEWEAQRDSHLGPHRSTPESRAAV 361

Qy    350 EEL-------EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLT 402
          :||       |: ||||||||||||||||||||||:|| |||||||:|||||||||||||
Db    362 QELSSSILAGEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLT 421

Qy    403 LLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
          |||||:|||||||||||||||:|||:|| ||:|||| || || ||
Db    422 LLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467


RESULT  12
US-08-670-964-4
; Sequence 4, Application US/08670964
; Patent No. 6010874
;  GENERAL INFORMATION:
;    APPLICANT:  Hardy, John A.
;    TITLE OF INVENTION:  EARLY ONSET ALZHEIMER'S DISEASE
;    TITLE OF INVENTION:  GENE AND GENE PRODUCTS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road - UW2220; P.O. Box 15
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/670,964
;      FILING DATE:  26-JUN-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/001,142
;      FILING DATE:  13-JUL-1995
;      APPLICATION NUMBER:  60/001,501
;      FILING DATE:  18-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Han, William T
;      REGISTRATION NUMBER:  34,344
;      REFERENCE/DOCKET NUMBER:  P50358
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5219
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  463 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-670-964-4


  Query Match             62.2%;  Score 1454;  DB 3;  Length 463;
  Best Local Similarity   65.3%;  Pred. No. 4.2e-141;
  Matches  303;  Conservative   40;  Mismatches   79;  Indels   42;  Gaps    8;

Qy     24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP---GRPPG------ 75
```

```
;     REFERENCE/DOCKET NUMBER:  P50361
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5219
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  463 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
US-08-670-479-18


  Query Match             61.5%;  Score 1437;  DB 2;  Length 463;
  Best Local Similarity   64.9%;  Pred. No. 2.3e-139;
  Matches  301;  Conservative   41;  Mismatches   80;  Indels   42;  Gaps    8;

Qy      24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP---GRPPG------ 75
           | | | |  :  |  ||   :    : :| |  | :  | :    |   ||| |
Db       3 ELPAPLSYFQNAQMSEDNHLS---NTNDNRERQEHNDRRSLGHPEPLSNGRPQGNSRQVV 59

Qy      75 -----LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTP 129
                :||||||||||||||||||||||||:|||||||| ||| |:||||||||||||
Db      60 EQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDTE 119

Qy     130 SVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLG 189
           :|||| |:|:||  |||||||||| |||||||||| || |||:|||:|||  ::||||
Db     120 TVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFEESFIYLG 179

Qy     190 EVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPE 249
           || ||||||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||||
Db     180 EVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYLPE 239

Qy     250 WSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDP 309
           |:||:||  |||||||||||||||||||||||:| :|||||||| ||| | ||: ||
Db     240 WTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERDETLFPALIYSSTMVWLVNMAEGDP 299

Qy     310 SSQGAL--QLPYDPE-MEEDSYDSFGE---PSYPEVFEPPLTGYPG-----------EE 351
            :|  :     |: |  | :| |:  |     : | :|     : |            :|
Db     300 EAQRRVSKNSKYNAESTERESQDTVAENDDGGFSEEWEAQRDSHLGPHRSTPESRAAVQE 359

Qy     352 L-------EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLL 404
           |       |: ||||||||||||||||||||:|| |||||||:|||||||||||||||
Db     360 LSSSILAGEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLL 419

Qy     405 LLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           |||:|||||||||||||:|||:|| ||:|||| || || ||
Db     420 LLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 463


RESULT  15
US-08-967-101-4
; Sequence 4, Application US/08967101
; Patent No. 5840540
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/967,101
;      FILING DATE:  10-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  467 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-967-101-4


  Query Match             60.7%;  Score 1418;  DB 2;  Length 467;
  Best Local Similarity   61.9%;  Pred. No. 2.1e-137;
  Matches  289;  Conservative   48;  Mismatches   86;  Indels   44;  Gaps    6;

Qy      24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP----GRPPG----- 75
           | | | |  :  |  ||  :::  ||| : :: ::  ||           |||
Db       3 EIPAPLSYFQNAQMSEDSHSSSAIRSQNDSQERQQQHDRQRLDNPEPISNGRPQSNSRQV 62

Qy      75 ------LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDT 128
                 :|||||||||||||||||||||||:||||||||| ||| |:||||||||||||
Db      63 VEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDT 122

Qy     129 PSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYL 188
           :|||| |:|:||  |||||||:||| ||||||||||| || |||:|||:||| |::|||
Db     123 ETVGQRALHSILNAAIMISVIVIMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYL 182

Qy     189 GEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLP 248
           ||| ||||| :|| |: | :||:| |||: ||||||| |||||||||||||||||||||
Db     183 GEVFKTYNVXVDYVTVALLIWNWGVVGMIAIHWKGPLRLQQAYLIMISALMALVFIKYLP 242

Qy     249 EWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLD 308
           ||:||:||  |||||||||||||||||||||||||||||| :|||||||| ||| | ||: |
Db     243 EWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGD 302

Qy     309 PSSQGALQLPYDPEME--------EDSYDSFGEPSYPEVFEPPLTGYPG----------- 350
           | :|   ::| :|:           :||      :  : | :|     : |
Db     303 PEAQ--RRVPKNPKYNTQRAERETQDSGSGNDDGGFSEEWEAQRDSHLGPHRSTPESRAA 360

Qy     350 -EEL-------EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCL 401
            :||       |: ||||||||||||||||||||:|| |||||||:|| |||||||||
Db     361 VQELSGSILTSEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACXVAILIGLCL 420

Qy     402 TLLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           ||||||::||  || ||||||| :| |:|| ||:|||| || || ||
Db     421 XLLLLAIYKKGXPAXPISITFGFVFXFATDYLVQPFMDQLAFHQFYI 467


Search completed: March 18, 2000, 19:55:31
Job time: 3229 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 18, 2000, 14:03:43 ; Search time 41.25 Seconds
(without alignments)
512.287 Million cell updates/sec

Title: US-08-509-359B-137
Perfect score: 2336
Sequence: 1 MLTFMASDSEEEVCDERTSL.....STDNLVRPFMDTLASHQLYI 448

Scoring table: BLOSUM62

Searched: 142080 seqs, 47169319 residues

Database : PIR_62:*

Word size : 0

Number of hits that pass the threshold : 142080
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2336	100.0	448	2	I58098	E5-1 protein - hum
2	2328	99.7	448	2	A56993	presenilin 2 - hum
3	2215.5	94.8	442	2	I39174	seven trans-membra
4	1624.5	69.5	449	2	JC5391	presenilin-beta -
5	1468	62.8	467	2	I78388	S182 protein - mou
6	1467	62.8	467	2	S58396	presenilin 1, spli
7	1454	62.2	463	2	S63683	presenilin I-463 -
8	1449.5	62.1	433	2	JC5390	presenilin-alpha -
9	1447	61.9	467	2	JC5080	presenilin 1 prote
10	1440	61.6	463	2	JC5081	presenilin 1 prote
11	1035.5	44.3	374	2	S63684	presenilin 1, spli
12	967	41.4	461	2	S60253	sel-12 protein - C
13	524.5	22.5	358	2	T15184	hypothetical prote
14	513	22.0	453	2	T00724	presenilin homolog
15	274	11.7	465	2	A43459	sperm membrane pro
16	119	5.1	531	2	T11596	hypothetical prote
17	110	4.7	2016	2	A38195	sodium channel pro
18	109	4.7	1840	1	CHRTM1	sodium channel pro
19	106	4.5	398	2	H75043	mg2+ transport pro
20	104	4.5	1476	1	A39901	cystic fibrosis tr
21	104	4.5	1476	1	A40303	cystic fibrosis tr
22	104	4.5	826	2	T02268	potassium transpor
23	103.5	4.4	1450	2	JC6139	cystic fibrosis tr
24	103	4.4	382	2	S47882	ubiquinol--cytochr
25	102.5	4.4	893	2	A47550	bride of sevenless
26	102	4.4	379	2	I48135	ubiquinol--cytochr
27	101.5	4.3	1681	2	A55138	sodium channel mNa
28	100.5	4.3	1951	2	S00320	sodium channel pro
29	100.5	4.3	1983	2	A60054	sodium channel pro
30	100.5	4.3	766	2	T01900	isp4 protein homol
31	100.5	4.3	324	2	S36646	integrin-associate
32	100	4.3	592	2	E70488	cytochrome-c oxida
33	99.5	4.3	379	2	JC6178	serotonin receptor
34	99	4.2	2019	2	A33996	sodium channel pro
35	99	4.2	521	2	T11710	probable amino aci
36	98	4.2	461	2	T11829	NADH dehydrogenase
37	98	4.2	1836	2	JS0648	sodium channel alp
38	98	4.2	1836	2	I51964	sodium channel alp
39	98	4.2	1836	2	I64893	sodium channel alp
40	98	4.2	1835	2	I54323	sodium channel alp
41	98	4.2	217	2	S01095	hypothetical prote
42	98	4.2	1695	2	JE0084	voltage-gated sodi
43	97.5	4.2	441	2	S13425	endothelin recepto
44	97	4.2	381	2	T11440	ubiquinol--cytochr
45	97	4.2	447	2	S52968	NADH dehydrogenase

ALIGNMENTS

RESULT 1
I58098
E5-1 protein - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C;Accession: I58098
R;Rogaev, E.I.; Sherrington, R.; Rogaeva, E.A.; Levesque, G.; Ikeda, M.; Liang, Y.; C
.; Cohen, D.; Lannfelt, L.; Fraser, P.E.; Rommens, J.M.; St George-Hyslop, P.H.
Nature 376, 775-778, 1995
A;Title: Familial Alzheimer's disease in kindreds with missense mutations in a gene o
A;Reference number: I58098; MUID:95379971
A;Accession: I58098
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-448 <RES>
A;Cross-references: GB:L44577; NID:g950347; PIDN:AAC42012.1; PID:g950348
C;Genetics:
A;Gene: E5-1
C;Superfamily: presenilin

Query Match 100.0%; Score 2336; DB 2; Length 448;
Best Local Similarity 100.0%; Pred. No. 8.1e-167;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60

Qy      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120

Qy     121 YTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 YTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180

Qy     181 FLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 FLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMA 240

Qy     241 LVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 LVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVW 300

Qy     301 TVGMAKLDPSSQGALQLPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGV 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TVGMAKLDPSSQGALQLPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGV 360

Qy     361 KLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISI 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISI 420

Qy     421 TFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           ||||||||||||||||||||||||||||
Db     421 TFGLIFYFSTDNLVRPFMDTLASHQLYI 448
```

```
Qy       7 SDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCS 66
           ||||:| |:|||||:::|||   | |:| |  |  | :     | :     ||
Db       5 SDSEDEECNERTSLITSESPPLPSYQDGVQASEGLETSYHRERQPDSTQNNED------- 58

Qy      67 GVP-GRPPGL-----------EEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYT 113
            || ||  |             |||||||||:||||||||||||:||||||||| |||
Db      58 -VPNGRTSGADAYNSETTVENEEEELTLKYGARHVIMLFVPVTLCMVVVVATIKSVSFYT 116

Qy     114 EKNGQLIYTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLI 173
           ||:||||||||:||| |||:|||||||||||||||:|||||||:|||||||||||||||
Db     117 EKDGQLIYTPFSEDTTSVGERLLNSVLNTLIMISVILVMTIFLVLLYKYRCYKFIHGWLI 176

Qy     174 MSSLMLLFLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLI 233
           :|||||||:|||||| || ||||:||||||| : :||||||||:|||||||| ||||||||
Db     177 LSSLMLLFMFTYIYLSEVFKTYNIAMDYPTLFMVIWNFGAVGMICIHWKGPLQLQQAYLI 236

Qy     234 MISALMALVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALI 293
           ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db     237 MISALMALVFIKYLPEWSAWVILGAISVYDLLAVLCPKGPLRMLVETAQERNEPIFPALI 296

Qy     294 YSSAMVWTVGMAKLDPSSQGAL--QLPY----DPEMEEDSYDSFGEPSYPEVFEPPLTGY 347
           |||||:||||||    :: | :  |: :     ||    : :   |           |
Db     297 YSSAMMWTVGMAD-SATADGRMNQQVQHIDRNTPEGANSTVEDAAETRIQ-------TQS 348

Qy     348 PGEELEEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLA 407
                : :|||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     349 NLSSEDPDEERGVKLGLGDFIFYSVLVGKAAATASGDWNTTLACFVAILIGLCLTLLLLA 408

Qy     408 VFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           |||||||||||||||||||||||:||||||||||||:||
Db     409 VFKKALPALPISITFGLIFYFSTDNIVRPFMDTLASHQMYI 449


RESULT   5
I78388
S182 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 29-Sep-1999
C;Accession: I78388
R;Sherrington, R.; Rogaev, E.I.; Liang, Y.; Rogaeva, E.A.; Levesque, G.; Ikeda, M.; Chi,
ero, I.; Pinessi, L.; Nee, L.; Chumakov, I.; Pollen, D.; Brookes, A.; Sanseau, P.; Polin
E.; Rommens, J.M.; St George-Hyslop, P.H.
Nature 375, 754-760, 1995
A;Title: Cloning of a gene bearing missense mutations in early-onset familial Alzheimer'
A;Reference number: I58095; MUID:95319502
A;Accession: I78388
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-467 <RES>
A;Cross-references: GB:L42177; NID:g904129; PIDN:AAC42094.1; PID:g904130
C;Superfamily: presenilin


  Query Match             62.8%;  Score 1468;  DB 2;  Length 467;
  Best Local Similarity   64.0%;  Pred. No. 3.9e-102;
  Matches  299;  Conservative   46;  Mismatches   78;  Indels   44;  Gaps    6;

Qy      24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP----GRPPG----- 75
           | | | |  :  |  ||  :::  ||| : :: ::  ||          |||
Db       3 EIPAPLSYFQNAQMSEDSHSSSAIRSQNDSQERQQQHDRQRLDNPEPISNGRPQSNSRQV 62

Qy      75 ------LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDT 128
                 :|||||||||||||||||||||:|||||||| ||| |:||||||||||||
Db      63 VEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDT 122

Qy     129 PSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYL 188
           :|||| |:|:||  |||||||:||| ||||||||||| || |||:|||:||| |::|||
Db     123 ETVGQRALHSILNAAIMISVIVIMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYL 182

Qy     189 GEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLP 248
           ||| |||||||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||
Db     183 GEVFKTYNVAVDYVTVALLIWNFGVVGMIAIHWKGPLRLQQAYLIMISALMALVFIKYLP 242

Qy     249 EWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLD 308
           ||:||:||  ||||||||||||||||||||||||||| :||||||| ||| | ||: |
Db     243 EWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGD 302

Qy     309 PSSQGALQLPYDPEME--------EDSYDSFGEPSYPEVFEPPLTGYPG----------- 350
           | :|   ::| :|:          :||     :  : | :|     : |
Db     303 PEAQ--RRVPKNPKYNTQRAERETQDSGSGNDDGGFSEEWEAQRDSHLGPHRSTPESRAA 360

Qy     350 -EEL-------EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCL 401
            :||       |: ||||||||||||||||||||:|| |||||||:||||||||||
Db     361 VQELSGSILTSEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCL 420

Qy     402 TLLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           ||||||:||||||||||||||:|||:|| ||:|||| || || ||
Db     421 TLLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467


RESULT   6
S58396
presenilin 1, splice form 467 - human
N;Alternate names: Alzheimer's disease protein 3; protein S182
C;Species: Homo sapiens (man)
C;Date: 29-Jan-1998 #sequence_revision 13-Feb-1998 #text_change 29-Sep-1999
C;Accession: S58396; S71401; S71402
R;Sherrington, R.; Rogaev, E.I.; Liang, Y.; Rogaeva, E.A.; Levesque, G.; Ikeda, M.; C
ero, I.; Pinessi, L.; Nee, L.; Chumakov, I.; Pollen, D.; Brookes, A.; Sanseau, P.; Po
E.; Rommens, J.M.; St George-Hyslop, P.H.
Nature 375, 754-760, 1995
A;Title: Cloning of a gene bearing missense mutations in early-onset familial Alzheim
A;Reference number: I58095; MUID:95319502
A;Accession: S58396
A;Molecule type: mRNA
A;Residues: 1-467 <SHE>
A;Cross-references: EMBL:L42110; NID:g904118; PIDN:AAB46416.1; PID:g904119
R;Vidal, R.; Ghiso, J.; Wisniewski, T.; Frangione, B.
FEBS Lett. 393, 19-23, 1996
A;Title: Alzheimer's presenilin 1 gene expression in platelets and megakaryocytes. Id
A;Reference number: S71401; MUID:96397521
A;Accession: S71401
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 24-32;254-256,290-292;316-317,376-379 <VIW>
A;Experimental source: Dami megakaryotic cell line (ATCC CRL-9792) and platelets
C;Genetics:
A;Gene: GDB:PSEN1; AD3; FAD; S182; PS1
A;Cross-references: GDB:135682; OMIM:104311
A;Map position: 14q24.3-14q24.3
C;Superfamily: presenilin
C;Keywords: alternative splicing; Alzheimer's disease; glycoprotein; transmembrane pr
F;82-100/Domain: transmembrane #status predicted <TM1>
F;133-154/Domain: transmembrane #status predicted <TM2>
F;164-185/Domain: transmembrane #status predicted <TM3>
F;195-213/Domain: transmembrane #status predicted <TM4>
F;221-238/Domain: transmembrane #status predicted <TM5>
F;244-264/Domain: transmembrane #status predicted <TM6>
F;281-301/Domain: transmembrane #status predicted <TM7>
F;408-428/Domain: transmembrane #status predicted <TM8>
F;433-453/Domain: transmembrane #status predicted <TM9>
F;279,405/Binding site: carbohydrate (Asn) (covalent) #status predicted


  Query Match             62.8%;  Score 1467;  DB 2;  Length 467;
  Best Local Similarity   65.7%;  Pred. No. 4.6e-102;
  Matches  306;  Conservative   39;  Mismatches   79;  Indels   42;  Gaps    8;

Qy      24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP-----GRPPG---- 75
           | | | |  :  |  ||  :   ||| :  : :|  ||    | |     ||| |
Db       3 ELPAPLSYFQNAQMSEDNHLSNTVRSQNDNRERQEHNDRR-SLGHPEPLSNGRPQGNSRQ 61

Qy      75 -------LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTED 127
                  :|||||||||||||||||||||:|||||||| ||| |:|||||||||
Db      62 VVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTED 121
```

presenilin 1 protein isoform 467 - lesser mouse lemur
C;Species: Microcebus murinus (lesser mouse lemur)
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 29-Sep-1999
C;Accession: JC5080
R;Calenda, A.; Mestre-Frances, N.; Czech, C.; Pradier, L.; Petter, A.; Bons, N.; Bellis,
Biochem. Biophys. Res. Commun. 228, 430-439, 1996
A;Title: Molecular cloning, sequencing, and brain expression of the presenilin 1 gene in
A;Reference number: JC5080; MUID:97079199
A;Accession: JC5080
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-467 <CAL>
A;Cross-references: EMBL:Z71333; NID:g1707591; PIDN:CAA95930.1; PID:e248527; PID:g170759
A;Experimental source: brain
C;Comment: This protein is an intergral membrane protein with seven transmembrane domain
C;Genetics:
A;Gene: ps1
A;Map position: 14
C;Superfamily: presenilin
C;Keywords: transmembrane protein
F;82-100/Domain: transmembrane #status predicted <TM1>
F;133-154/Domain: transmembrane #status predicted <TM2>
F;164-185/Domain: transmembrane #status predicted <TM3>
F;195-213/Domain: transmembrane #status predicted <TM4>
F;221-238/Domain: transmembrane #status predicted <TM5>
F;244-261/Domain: transmembrane #status predicted <TM6>
F;408-428/Domain: transmembrane #status predicted <TM7>

Query Match 61.9%; Score 1447; DB 2; Length 467;
Best Local Similarity 64.4%; Pred. No. 1.4e-100;
Matches 300; Conservative 41; Mismatches 83; Indels 42; Gaps 7;

```
Qy      24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP-----GRPPG---- 75
           | | | |  :  |  ||   :   ||| :  : ::  ||    | |     ||| |
Db       3 ELPAPLSYFQNAQMSEDNHLSNTVRSQNDNREQQDHGDRRRL-GNPEPLSNGRPQGNSGP 61

Qy      75 -------LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTED 127
                 :||||||||||||||||||||||:||||||||| ||| |:||||||||||
Db      62 VVERDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTED 121

Qy     128 TPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIY 187
           | :|||| |:|||| ||||||||||| |||||||||| || |||:|||:||| |::||
Db     122 TETVGQRALHSVLNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIY 181

Qy     188 LGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYL 247
           |||| ||||||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||
Db     182 LGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYL 241

Qy     248 PEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKL 307
           |||:||:||  |||||||||||||||||||||||||||| :|||||||| ||| | ||:
Db     242 PEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEG 301

Qy     308 DPSSQGAL--QLPYD------------PEMEEDSYDSFGEPSYPEVFEP----PLTGYPG 349
           || :|  :    |:             || ::  :    |        |     ::
Db     302 DPEAQRRVSKNTKYNAQGTEREAQASVPENDDGGFSEEWEAQRDSQLGPHRSTSVSRAAV 361

Qy     350 EEL-------EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLT 402
           :|:       |: |||||||||||:||||||||:|| |||||||:|||||||||||||
Db     362 QEISSSIPASEDPEERGVKLGLGDFVFYSVLVGKASATASGDWNTTIACFVAILIGLCLT 421

Qy     403 LLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           |||||:||||||||||||||:|||:|| ||:|||| || || ||
Db     422 LLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467
```

RESULT 10
JC5081
presenilin 1 protein isoform 463 - lesser mouse lemur
C;Species: Microcebus murinus (lesser mouse lemur)
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 13-Sep-1998
C;Accession: JC5081
R;Calenda, A.; Mestre-Frances, N.; Czech, C.; Pradier, L.; Petter, A.; Bons, N.; Bellis,
Biochem. Biophys. Res. Commun. 228, 430-439, 1996
A;Title: Molecular cloning, sequencing, and brain expression of the presenilin 1 gene
A;Reference number: JC5080; MUID:97079199
A;Contents: brain
A;Accession: JC5081
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-463 <CAL>
A;Cross-references: EMBL:Z71333
C;Comment: This protein is an intermembrane protein with seven transmembrane domains.
C;Genetics:
A;Gene: ps1
A;Map position: 14
C;Superfamily: presenilin
C;Keywords: transmembrane protein
F;78-96/Domain: transmembrane #status predicted <TM1>
F;129-150/Domain: transmembrane #status predicted <TM2>
F;160-181/Domain: transmembrane #status predicted <TM3>
F;191-209/Domain: transmembrane #status predicted <TM4>
F;217-234/Domain: transmembrane #status predicted <TM5>
F;240-257/Domain: transmembrane #status predicted <TM6>
F;404-424/Domain: transmembrane #status predicted <TM7>

Query Match 61.6%; Score 1440; DB 2; Length 463;
Best Local Similarity 64.2%; Pred. No. 4.7e-100;
Matches 300; Conservative 39; Mismatches 80; Indels 48; Gaps 8;

```
Qy      24 ESPTPRSCQEGRQGPEDG--ENTAQWRSQENEED----GEEDPDRYVCSGVPGRPPG--- 75
           | | | |  :  |  ||    ||   | |::  |    |  :|         ||| |
Db       3 ELPAPLSYFQNAQMSEDNHLSNTNDNREQQDHGDRRRLGNPEPLS------NGRPQGNSG 56

Qy      75 --------LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTE 126
                  :|||||||||||||||||||||:|||||||| ||| |:||||||||||
Db      57 PVVERDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTE 116

Qy     127 DTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYI 186
           || :|||| |:|||| ||||||||||| |||||||||| || |||:|||:||| |::|
Db     117 DTETVGQRALHSVLNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFI 176

Qy     187 YLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKY 246
           ||||| ||||||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||
Db     177 YLGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKY 236

Qy     247 LPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAK 306
           ||||:||:||  |||||||||||||||||||||||||||| :|||||||| ||| | ||:
Db     237 LPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAE 296

Qy     307 LDPSSQGAL--QLPYD------------PEMEEDSYDSFGEPSYPEVFEP----PLTGYP 348
           || :|  :    |:             || ::  :    |        |     ::
Db     297 GDPEAQRRVSKNTKYNAQGTEREAQASVPENDDGGFSEEWEAQRDSQLGPHRSTSVSRAA 356

Qy     349 GEEL-------EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCL 401
           :|:       |: |||||||||||:||||||||:|| |||||||:|||||||||||||
Db     357 VQEISSSIPASEDPEERGVKLGLGDFVFYSVLVGKASATASGDWNTTIACFVAILIGLCL 416

Qy     402 TLLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           ||||||:||||||||||||||:|||:|| ||:|||| || || ||
Db     417 TLLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 463
```

RESULT 11
S63684
presenilin 1, splice form 374 - human
N;Alternate names: Alzheimer's disease protein 3
C;Species: Homo sapiens (man)
C;Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 29-Sep-1999
C;Accession: S63684
R;Sahara, N.; Yahagi, Y.; Takagi, H.; Kondo, T.; Okochi, M.; Usami, M.; Shirasawa, T.
FEBS Lett. 381, 7-11, 1996
A;Title: Identification and characterization of presenilin I-467, I-463 and I-374.
A;Reference number: S63683; MUID:96193901
A;Accession: S63684

```
          ||||:   : || |   |  : |
Db    336 FYFSSHIALTPFTDLCTSQLILI 358


RESULT  14
T00724
presenilin homolog F22O13.19 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 20-Sep-1999
C;Accession: T00724
R;Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Con
eologis, A.; Ecker, J.R.
submitted to the EMBL Data Library, April 1998
A;Description: Genomic sequence for Arabidopsis thaliana BAC F22O13.
A;Reference number: Z14200
A;Accession: T00724
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-453 <SHI>
A;Cross-references: EMBL:AC003981; NID:g3063438; PID:g3063457; GSPDB:GN0059; ATSP:F22O13
C;Genetics:
A;Gene: ATSP:F22O13.19
A;Map position: 1
A;Introns: 108/1
C;Superfamily: presenilin


  Query Match             22.0%;  Score 513;  DB 2;  Length 453;
  Best Local Similarity   31.1%;  Pred. No. 5.2e-31;
  Matches  136;  Conservative   78;  Mismatches  130;  Indels   94;  Gaps   12;

Qy      88 VIMLFVPVTLCMIVVVATIKSVRFYTEKN----GQLIYTPFTEDTPSVGQRLLNSVLNTL 143
           :| :  ||::|| :||    |:   ::        |||     |: :|  :|  |: | :
Db      13 IIGVMAPVSICMFLVVLLTYSLSVTSDPQIRSAANLIYIENPSDSTTV--KLEGSLANAI 70

Qy     144 IMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYNVAMDYPT 203
           : : :|  :|  ||:|: |    |:  ::  |:  :|          ::: ::: :|  |
Db      71 VFVVLIAAVTFILVLLFYYNFTNFLKHYMRFSAFFVLGTMGGAIFLSIIQHFSIPVDSIT 130

Qy     204 LLLTVWNFGAVGMVCIHWKG-PLVLQQAYLIMISALMALVFIKYLPEWSAWVILGAISVY 262
             : ::||  :| : :   | |:||:| |::::  ::|  | | ||||: | || |:::|
Db     131 CFILLFNFTILGTLSVFAGGIPIVLRQCYMVVMGIVVAAWFTK-LPEWTTWFILVALALY 189

Qy     263 DLVAVLCPKGPLRMLVETAQERNEPIFPALIY--------------SSAMVWTVGMAKL 307
           |||||| | |||::||| |  |:| : ||::|               |::   ||   :
Db     190 DLVAVLAPGGPLKLLVELASSRDEEL-PAMVYEARPTVSSGNQRRNRGSSLRALVGGGGV 248

Qy     308 DPSSQGALQLPYDPEM----EEDSY---------------DSFGEPSYPEVFEPPLTGY 347
             |    ||   : ::     |:|:                |  |  |   :   || |
Db     249 SDSGSVELQAVRNHDVNQLGRENSHNMDYNAIAVRDIDNVDDGIGNGSRGGLERSPLVGS 308

Qy     348 PG-----------------------------------EE---LEEEE------ERGV 360
           |                                   ||   ||| :        ||:
Db     309 PSASEHSTSVGTRGNMEDRESVMDEEMSPLVELMGWGDNREEARGLEESDNVVDISNRGI 368

Qy     361 KLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISI 420
           |||||||||||||||:||     |  |  ||::||: ||  ||:||:|: :|||||||||
Db     369 KLGLGDFIFYSVLVGRAAMY---DLMTVYACYLAIISGLGCTLILLSVYNRALPALPISI 425

Qy     421 TFGLIFYFSTDNLVRPFM 438
             |::||| |  |: ||:
Db     426 MLGVVFYFLTRLLMEPFV 443


RESULT  15
A43459
sperm membrane protein spe-4 - Caenorhabditis elegans
N;Alternate names: probable integral membrane protein
C;Species: Caenorhabditis elegans
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 09-Sep-1997
C;Accession: A43459; S24632; S24633
R;L'Hernault, S.W.; Arduengo, P.M.
J. Cell Biol. 119, 55-68, 1992
A;Title: Mutation of a putative sperm membrane protein in Caenorhabditis elegans prev
A;Reference number: A43459; MUID:92407040
A;Accession: A43459
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA; mRNA
A;Residues: 1-465 <LHE>
A;Cross-references: EMBL:Z14067; NID:g6868; PID:g6869; EMBL:Z14066; NID:g6870; PID:g6
A;Experimental source: strain Bristol N2
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1992
C;Genetics:
A;Introns: 69/3; 154/3; 200/1; 224/3; 300/1; 386/1; 435/1


  Query Match             11.7%;  Score 274;  DB 2;  Length 465;
  Best Local Similarity   21.5%;  Pred. No. 3.4e-13;
  Matches   93;  Conservative   80;  Mismatches  127;  Indels  132;  Gaps   14;

Qy     114 EKNGQLIYTPFTEDT--PSVGQRLLNSVLN----TLIMISVIVVMTIFLVVLYKYRCYKF 167
           | | :|  | | : :    : |  ||:  :|     |::  |  :|  |  ||: :|  :
Db      42 EVNSELSKTYFLDPSFEQTTGNLLLDGFINGVGTILVLGCVSFIMLAF--VLFDFR--RI 97

Qy     168 IHGWLIMSSLMLLF---------LFTYIYLGEVLKTYNVAMDYPTLLL-----TVWNFGA 213
           :  || :| |::||          :|: ::  :     |      | |::|      |: ||
Db      98 VKAWLTLSCLLILFGVSAQTLHDMFSQVFDQDDNNQY-----YMTIVLIVVPTVVYGFG- 152

Qy     214 VGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVILGAISVYDLVAVLCPKGP 273
           |:        |:| | :::   :|::: :::  |  : | :|  :  :|| ||| | ||
Db     152 -GIYAFFSNSSLILHQIFVVTNCSLISVFYLRVFPSKTTWFVLWIVLFWDLFAVLAPMGP 210

Qy     274 LRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQG----------ALQLPYDPE 322
           |: : | | : :: :   :::|:         :  : :::|             ::     |
Db     211 LKKVQEKASDYSKCVLNLIMFSANEKRLTAGSNQEETNEGEESTIRRTVKQTIEYYTKRE 270

Qy     323 MEEDSY-------------DSF---------GEPSYPEVFEPPLTGYPGEELEEEEE--- 358
           ::| :               ||          |||  |: |   |    ||| ::|
Db     271 AQDDEFYQKIRQRRAAINPDSVPTEHSPLVEAEPSPIELKEKNST----EELSDDESDTS 326

Qy     358 ----------------------------------------------------------R 358

Db     327 ETSSGSSNLSSSDSSTTVSTSDISTAEECDQKEWDDLVSNSLPNNDKRPATAADALNDGE 386

Qy     359 GVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPI 418
           ::|| |||:|||:|:|:|||:|       ::  : || || :|| : :  :   ||||:
Db     387 VLRLGFGDFVFYSLLIGQAAASGCP--FAVISAALGILFGLVVTLTVFSTEESTTPALPL 444

Qy     419 SITFGLIFYFST 430
           :  |   |||:
Db     445 PVICGTFCYFSS 456


Search completed: March 18, 2000, 14:11:52
Job time: 489 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 18, 2000, 16:08:11 ; Search time 32.57 Seconds
(without alignments)
410.791 Million cell updates/sec

Title: US-08-509-359B-137
Perfect score: 2336
Sequence: 1 MLTFMASDSEEEVCDERTSL.....STDNLVRPFMDTLASHQLYI 448

Scoring table: BLOSUM62

Searched: 82229 seqs, 29864866 residues

Database : SwissProt_38:*

Word size : 0

Number of hits that pass the threshold : 82229

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	2336	100.0	448	1	PSN2_HUMAN	P49810	homo sapien
2	2232	95.5	448	1	PSN2_MOUSE	Q61144	mus musculu
3	2227	95.3	448	1	PSN2_RAT	O88777	rattus norv
4	2218	94.9	445	1	PSN2_MICMU	P79801	microcebus
5	1624.5	69.5	449	1	PSN2_XENLA	O12977	xenopus lae
6	1468	62.8	467	1	PSN1_MOUSE	P49769	mus musculu
7	1467	62.8	467	1	PSN1_HUMAN	P49768	homo sapien
8	1459.5	62.5	468	1	PSN1_RAT	P97887	rattus norv
9	1449.5	62.1	433	1	PSN1_XENLA	O12976	xenopus lae
10	1447	61.9	467	1	PSN1_MICMU	P79802	microcebus
11	1150	49.2	541	1	PSN_DROME	O02194	drosophila
12	1043	44.6	836	1	YL4K_CAEEL	Q20076	caenorhabdi
13	983	42.1	461	1	SE12_CAEEL	P52166	caenorhabdi
14	524.5	22.5	358	1	HOP1_CAEEL	O02100	caenorhabdi
15	513	22.0	453	1	PSNH_ARATH	O64668	arabidopsis
16	274	11.7	465	1	SPE4_CAEEL	Q01608	caenorhabdi
17	119	5.1	531	1	YDFG_SCHPO	Q10487	schizosacch
18	110	4.7	2016	1	CIN5_HUMAN	Q14524	homo sapien
19	109	4.7	1840	1	CIN4_RAT	P15390	rattus norv
20	108	4.6	381	1	CYB_DASCR	Q34302	dasycercus
21	104	4.5	1476	1	CFTR_MOUSE	P26361	mus musculu
22	103.5	4.4	1450	1	CFTR_RABIT	Q00554	oryctolagus
23	103	4.4	382	1	CYB_DIDMA	P41303	didelphis m
24	103	4.4	380	1	CYB_MICLO	P56731	microtus lo
25	102.5	4.4	893	1	BOSS_DROVI	Q24738	drosophila
26	102.5	4.4	381	1	CYB_ANTFL	Q33706	antechinus
27	102	4.4	381	1	CYB_DASMA	O03522	dasyurus ma
28	102	4.4	381	1	CYB_PSENI	Q35553	pseudantech
29	101.5	4.3	381	1	CYB_NINYV	Q35196	ningaui yvo
30	101.5	4.3	381	1	CYB_PARAP	Q35377	parantechin
31	101	4.3	381	1	CYB_PLAMS	Q35533	planigale m
32	100.5	4.3	1951	1	CIN3_RAT	P08104	rattus norv
33	100.5	4.3	381	1	CYB_PSEMD	O03543	pseudantech
34	100	4.3	381	1	CYB_DASGE	O20604	dasyurus ge
35	99	4.2	2019	1	CIN5_RAT	P15389	rattus norv
36	99	4.2	381	1	CYB_ANTMI	O63534	antechinus
37	99	4.2	381	1	CYB_PHATA	Q35673	phascogale
38	98.5	4.2	872	1	CIQ3_HUMAN	O43525	homo sapien
39	98	4.2	1836	1	CIN4_HUMAN	P35499	homo sapien
40	98	4.2	381	1	CYB_ANTME	Q33782	antechinus
41	98	4.2	381	1	CYB_DASAL	Q34289	dasyurus al
42	98	4.2	381	1	CYB_SMICR	Q35810	sminthopsis
43	98	4.2	460	1	NU4M_GADMO	P55781	gadus morhu
44	98	4.2	217	1	YPRA_ECOLI	P13974	escherichia
45	97.5	4.2	381	1	CYB_DASHA	Q34321	dasyurus ha

ALIGNMENTS

RESULT 1
PSN2_HUMAN
ID PSN2_HUMAN STANDARD; PRT; 448 AA.
AC P49810;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PRESENILIN 2 (PS-2) (STM-2) (E5-1) (AD3LP) (AD5).
GN PSEN2 OR PSNL2 OR AD4 OR PS2 OR STM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT FAD ILE-141.
RX MEDLINE; 95365816.
RA LEVY-LAHAD E., WASCO W., POORKAJ P., ROMANO D.M., OSHIMA J.,
RA PETTINGELL W.H., YU C.-E., JONDRO P.D., SCHMIDT S.D., WANG K.,
RA CROWLEY A.C., FU Y.-H., GUENETTE S.Y., GALAS D., NEMENS E.,
RA WIJSMAN E.M., BIRD T.D., SCHELLENBERG G.D., TANZI R.E.;
RT "Candidate gene for the chromosome 1 familial Alzheimer's disease
RT locus.";
RL Science 269:973-977(1995).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS FAD ILE-141 AND VAL-239.
RC TISSUE=BRAIN, AND COLON;
RX MEDLINE; 95379971.
RA ROGAEV E.I., SHERRINGTON R., ROGAEVA E.A., LEVESQUE G., IKEDA M.,
RA LIANG Y., CHI H., LIN C., HOLMAN K., TSUDA T., MAR L., SORBI S.,
RA NACMIAS B., PIACENTINI S., AMADUCCI L., CHUMAKOV I., COHEN D.,
RA LANNFELT L., FRASER P.E., ROMMENS J.M., ST GEORGE-HYSLOP P.H.;
RT "Familial Alzheimer's disease in kindreds with missense mutations in
RT a gene on chromosome 1 related to the Alzheimer's disease type 3
RT gene.";
RL Nature 376:775-778(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96109229.
RA LI J., MA J., POTTER H.;
RT "Identification and expression analysis of a potential familial
RT Alzheimer disease gene on chromosome 1 related to AD3.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:12180-12184(1995).
RN [4]
RP SEQUENCE FROM N.A.
RA LEVY-LAHAD E., POORKAJ P., WANG K., FU Y.H., OSHIMA J.,
RA MULLIGAN J., SCHELLENBERG G.D.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP REVIEW ON VARIANTS.
RX MEDLINE; 98180715.
RA CRUTS M., VAN BROECKHOVEN C.;
RT "Presenilin mutations in Alzheimer's disease.";
RL Hum. Mutat. 11:183-190(1998).
RN [6]
RP VARIANT AD HIS-62.
RX MEDLINE; 98046005.
RA CRUTS M., VAN DUIJN C.M., BACKHOVENS H., VAN DEN BROECK M.,
RA WEHNERT A., SERNEELS S., SHERRINGTON R., HUTTON M., HARDY J.,
RA ST GEORGE-HYSLOP P.H., HOFMAN A., VAN BROECKHOVEN C.;
RT "Estimation of the genetic contribution of presenilin-1 and -2
RT mutations in a population-based study of presenile Alzheimer
RT disease.";
RL Hum. Mol. Genet. 7:43-51(1998).

```
DR   EMBL; U57325; AAC53311.1; -.
DR   EMBL; AF038935; AAB92660.1; -.
DR   EMBL; U49111; AAC52935.1; -.
DR   MGD; MGI:109284; PSEN2.
DR   PFAM; PF01080; Presenilin; 1.
KW   Transmembrane; Alternative initiation.
FT   CHAIN         1    448       PRESENILIN 2.
FT   CHAIN       298    448       PRESENILIN 2-SHORT.
FT   TRANSMEM     88    106       POTENTIAL.
FT   TRANSMEM    142    160       POTENTIAL.
FT   TRANSMEM    167    188       POTENTIAL.
FT   TRANSMEM    203    219       POTENTIAL.
FT   TRANSMEM    230    246       POTENTIAL.
FT   TRANSMEM    253    269       POTENTIAL.
FT   TRANSMEM    288    305       POTENTIAL.
FT   TRANSMEM    387    406       POTENTIAL.
FT   TRANSMEM    413    429       POTENTIAL.
FT   CONFLICT     87     87       R -> H (IN REF. 2).
FT   CONFLICT    226    226       A -> V (IN REF. 2).
FT   CONFLICT    324    324       MISSING (IN REF. 2).
SQ   SEQUENCE   448 AA;  49955 MW;  680ACF19 CRC32;


  Query Match             95.5%;  Score 2232;  DB 1;  Length 448;
  Best Local Similarity   95.5%;  Pred. No. 7.3e-147;
  Matches  428;  Conservative    6;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60
           || ||||||||||||||||||||||| ||||||| |||||:|||||:||:||| ||||
Db       1 MLAFMASDSEEEVCDERTSLMSAESPTSRSCQEGRPGPEDGESTAQWRTQESEEDCEEDP 60

Qy      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120
           ||| ||| |||| |||||||||||| ||||||||||||||||||||||||||||||||||
Db      61 DRYACSGAPGRPSGLEEELTLKYGAKRVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120

Qy     121 YTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 YTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180

Qy     181 FLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMA 240
           |||||||||||||||||||||| | |||||||||||||||| |||||||:||||||
Db     181 FLFTYIYLGEVLKTYNVAMDYPTLFLAVWNFGAVGMVCIHWKGPLALQQAYLIVISALMA 240

Qy     241 LVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 LVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVW 300

Qy     301 TVGMAKLDPSSQGALQLPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERGV 360
           |||||||||||||||||||||||||||||||||| || || ||||||||||||||||
Db     301 TVGMAKLDPSSQGALQLPYDPEMEEDSYDSFGEPSYPEAFEAPLPGYPGEELEEEEERGV 360

Qy     361 KLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISI 420
           ||||||||||||||||||:||||||||||:||||||||||||||||||||||||||||||
Db     361 KLGLGDFIFYSVLVGKAAATGNGDWNTTLACFIAILIGLCLTLLLLAVFKKALPALPISI 420

Qy     421 TFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           ||||||||||||||||||||||||||||
Db     421 TFGLIFYFSTDNLVRPFMDTLASHQLYI 448


RESULT   3
PSN2_RAT
ID   PSN2_RAT       STANDARD;      PRT;   448 AA.
AC   O88777; O35546; O08947;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   PRESENILIN 2 (PS-2).
GN   PSEN2 OR PSNL2 OR PS2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=BRAIN;
RA   FRENTZEL S., ABDEL A.S., LUEBBERT H.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=BRAIN;
RX   MEDLINE; 97473536.
RA   TAKAHASHI H., MERCKEN M., NAKAZATO Y., NOGUCHI K., MURAYAMA M.,
RA   IMAHORI K., TAKASHIMA A.;
RT   "Cloning of cDNA and expression of the gene encoding rat
RT   presenilin-2.";
RL   Gene 197:383-387(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=BRAIN;
RX   MEDLINE; 98207716.
RA   TANAHASHI H., TABIRA T.;
RT   "Cloning of the cDNA encoding rat presenilin-2.";
RL   Biochim. Biophys. Acta 1396:259-262(1998).
CC   -!- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE
CC       EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE. MAY
CC       FUNCTION IN THE CYTOPLASMIC PARTITIONING OF PROTEINS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X99267; CAA67663.1; -.
DR   EMBL; D83700; BAA22832.1; -.
DR   EMBL; AB004454; BAA20406.1; -.
DR   PFAM; PF01080; Presenilin; 1.
KW   Transmembrane.
FT   TRANSMEM     88    106       POTENTIAL.
FT   TRANSMEM    142    160       POTENTIAL.
FT   TRANSMEM    167    188       POTENTIAL.
FT   TRANSMEM    203    219       POTENTIAL.
FT   TRANSMEM    230    246       POTENTIAL.
FT   TRANSMEM    253    269       POTENTIAL.
FT   TRANSMEM    288    305       POTENTIAL.
FT   TRANSMEM    387    406       POTENTIAL.
FT   TRANSMEM    413    429       POTENTIAL.
FT   CONFLICT      7      7       S -> T (IN REF. 1).
FT   CONFLICT     86     87       KH -> ND (IN REF. 3).
SQ   SEQUENCE   448 AA;  50051 MW;  EE0DF681 CRC32;


  Query Match             95.3%;  Score 2227;  DB 1;  Length 448;
  Best Local Similarity   95.3%;  Pred. No. 1.6e-146;
  Matches  427;  Conservative    6;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDP 60
           |||||||||||||||||||||||| |||| | ||||||||||||||||:| ||||
Db       1 MLTFMASDSEEEVCDERTSLMSAESPTSRSCQDSRPGPEDGENTAQWRSQENEDDCEEDP 60

Qy      61 DRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120
           | | |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db      61 DHYACSGVPGRPSGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLI 120

Qy     121 YTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 YTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLL 180

Qy     181 FLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMA 240
           |||||||||| |||||||||| | |||||||||||||||||||||||||:||||||
Db     181 FLFTYIYLGEVFKTYNVAMDYPTLFLAVWNFGAVGMVCIHWKGPLVLQQAYLIVISALMA 240
```

```
FT   TRANSMEM    415    435       POTENTIAL.
FT   CARBOHYD    387    387       POTENTIAL.
SQ   SEQUENCE   449 AA;  50274 MW;  DF6EAF4E CRC32;


  Query Match             69.5%;  Score 1624.5;  DB 1;  Length 449;
  Best Local Similarity   72.0%;  Pred. No. 4.9e-105;
  Matches  332;  Conservative   34;  Mismatches   60;  Indels   35;  Gaps    7;

Qy       7 SDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCS 66
           ||||:| |:|||||:::|||   | |:| |  |  | :     | :     ||
Db       5 SDSDEDECNERTSLITSESPPLPSYQDGVQASEGLETSYHRERQPDSTQNNED------- 58

Qy      67 GVP-GRPPGL-----------EEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYT 113
            || ||  |             ||||||||||:||||||||||||:||||||||| |||
Db      58 -VPNGRTSGADAYNSETTVENEEEELTLKYGARHVIMLFVPVTLCMVVVVATIKSVSFYT 116

Qy     114 EKNGQLIYTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLI 173
           ||:|||||||:||| |||:|||||||||||||||:|||||||:|||||||||||||||
Db     117 EKDGQLIYTPFSEDTTSVGERLLNSVLNTLIMISVILVMTIFLVLLYKYRCYKFIHGWLI 176

Qy     174 MSSLMLLFLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLI 233
           :|||||||:|||||| || ||||:|||||| : :|||||||:|||||||| |||||||
Db     177 LSSLMLLFMFTYIYLSEVFKTYNIAMDYPTLFMVIWNFGAVGMICIHWKGPLQLQQAYLI 236

Qy     234 MISALMALVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALI 293
           ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db     237 MISALMALVFIKYLPEWSAWVILGAISVYDLLAVLCPKGPLRMLVETAQERNEPIFPALI 296

Qy     294 YSSAMVWTVGMAKLDPSSQGAL--QLPY----DPEMEEDSYDSFGEPSYPEVFEPPLTGY 347
           |||||:||||||    :: | :  |: :     ||    : :   |           |
Db     297 YSSAMMWTVGMAD-SATADGRMNQQVQHIDRNTPEGANSTVEDAAETRIQ-------TQS 348

Qy     348 PGEELEEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLA 407
           : :|||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     349 NLSSEDPDEERGVKLGLGDFIFYSVLVGKAAATASGDWNTTLACFVAILIGLCLTLLLLA 408

Qy     408 VFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           ||||||||||||||||||||||||:||||||||||||:||
Db     409 VFKKALPALPISITFGLIFYFSTDNIVRPFMDTLASHQMYI 449


RESULT   6
PSN1_MOUSE
ID   PSN1_MOUSE     STANDARD;      PRT;   467 AA.
AC   P49769;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-DEC-1999 (Rel. 39, Last annotation update)
DE   PRESENILIN 1 (PS-1) (S182 PROTEIN).
GN   PSEN1 OR PSNL1 OR AD3H.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 95319502.
RA   SHERRINGTON R., ROGAEV E.I., LIANG Y., ROGAEVA E.A., LEVESQUE G.,
RA   IKEDA M., CHI H., LIN C., LI G., HOLMAN K., TSUDA T., MAR L.,
RA   FONCIN J.-F., BRUNI A.C., MONTESI M.P., SORBI S., RAINERO I.,
RA   PINESSI L., NEE L., CHUMAKOV I., POLLEN D., BROOKES A.,
RA   SANSEAU P., POLINSKY R.J., WASCO W., DA SILVA H.A.R., HAINES J.L.,
RA   PERICAK-VANCE M.A., TANZI R.E., ROSES A.D., FRASER P.E.,
RA   ROMMENS J.M., ST GEORGE-HYSLOP P.H.;
RT   "Cloning of a gene bearing missense mutations in early-onset familial
RT   Alzheimer's disease.";
RL   Nature 375:754-760(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SVJ;
RX   MEDLINE; 97442406.
RA   MITSUDA N., ROSES A.D., VITEK M.P.;
RT   "Transcriptional regulation of the mouse presenilin-1 gene.";
RL   J. Biol. Chem. 272:23489-23497(1997).
CC   -!- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE
CC       EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE.
CC       MAY FUNCTION IN THE CYTOPLASMIC PARTITIONING OF PROTEINS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L42177; AAC42094.1; -.
DR   EMBL; AF007560; AAB72049.1; -.
DR   MGD; MGI:1202717; PSEN1.
DR   PFAM; PF01080; Presenilin; 1.
KW   Transmembrane; Glycoprotein.
FT   TRANSMEM     83    103       POTENTIAL.
FT   TRANSMEM    133    153       POTENTIAL.
FT   TRANSMEM    161    181       POTENTIAL.
FT   TRANSMEM    195    215       POTENTIAL.
FT   TRANSMEM    221    241       POTENTIAL.
FT   TRANSMEM    244    264       POTENTIAL.
FT   TRANSMEM    281    301       POTENTIAL.
FT   TRANSMEM    408    428       POTENTIAL.
FT   TRANSMEM    433    453       POTENTIAL.
FT   CARBOHYD     30     30       POTENTIAL.
FT   CARBOHYD    279    279       POTENTIAL.
FT   CARBOHYD    405    405       POTENTIAL.
SQ   SEQUENCE   467 AA;  52639 MW;  3AE1350D CRC32;


  Query Match             62.8%;  Score 1468;  DB 1;  Length 467;
  Best Local Similarity   64.0%;  Pred. No. 3e-94;
  Matches  299;  Conservative   46;  Mismatches   78;  Indels   44;  Gaps    6;

Qy      24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP----GRPPG----- 75
           | | | |  :  |  ||  :::  ||| : :: ::  ||           |||
Db       3 EIPAPLSYFQNAQMSEDSHSSSAIRSQNDSQERQQQHDRQRLDNPEPISNGRPQSNSRQV 62

Qy      75 ------LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTEDT 128
                 :|||||||||||||||||||||:|||||||| ||| |:||||||||||||
Db      63 VEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTEDT 122

Qy     129 PSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYL 188
           :|||| |:|:|| ||||||:||| ||||||||||| || |||:|||:||| |::|||
Db     123 ETVGQRALHSILNAAIMISVIVIMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIYL 182

Qy     189 GEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLP 248
           ||| ||||||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||||
Db     183 GEVFKTYNVADVYVTALLIWNFGVVGMIAIHWKGPLRLQQAYLIMISALMALVFIKYLP 242

Qy     249 EWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLD 308
           ||:||:||  |||||||||||||||||||||||||||| :|||||||| ||| | ||: |
Db     243 EWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEGD 302

Qy     309 PSSQGALQLPYDPEME--------EDSYDSFGEPSYPEVFEPPLTGYPG----------- 350
           | :|   ::| :|:          :||      :  : | :|     : |
Db     303 PEAQ--RRVPKNPKYNTQRAERETQDSGSGNDDGGFSEEWEAQRDSHLGPHRSTPESRAA 360

Qy     350 -EEL-------EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCL 401
            :||      |: ||||||||||||||||||:|| |||||||:||||||||||
Db     361 VQELSGSILTSEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCL 420

Qy     402 TLLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           ||||||:||||||||||||:|||:|| ||:|||| || || ||
Db     421 TLLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467
```

```
RA   ROYTTA M., LILIUS L., EEROLA A., ST GEORGE-HYSLOP P.H., FREY H.,
RA   LANNFELT L.;
RT   "The Glu318Gly mutation of the presenilin-1 gene does not necessarily
RT   cause Alzheimer's disease.";
RL   Ann. Neurol. 44:965-967(1998).
RN   [15]
RP   VARIANT GLY-318.
RX   MEDLINE; 99066775.
RA   ALDUDO J., BULLIDO M.J., FRANK A., VALDIVIESO F.;
RT   "Missense mutation E318G of the presenilin-1 gene appears to be a
RT   nonpathogenic polymorphism.";
RL   Ann. Neurol. 44:985-986(1998).
RN   [16]
RP   VARIANTS AD VAL-79; CYS-115; VAL-231, AND VARIANT GLY-318.
RX   MEDLINE; 98046005.
RA   CRUTS M., VAN DUIJN C.M., BACKHOVENS H., VAN DEN BROECK M.,
RA   WEHNERT A., SERNEELS S., SHERRINGTON R., HUTTON M., HARDY J.,
RA   ST GEORGE-HYSLOP P.H., HOFMAN A., VAN BROECKHOVEN C.;
RT   "Estimation of the genetic contribution of presenilin-1 and -2
RT   mutations in a population-based study of presenile Alzheimer
RT   disease.";
RL   Hum. Mol. Genet. 7:43-51(1998).
RN   [17]
RP   VARIANTS AD D-120; R-163; V-209; V-260; L-264; Y-410 AND P-426.
RX   MEDLINE; 98180720.
RA   POORKAJ P., SHARMA V., ANDERSON L., NEMENS E., ALONSO M.E., ORR H.,
RA   WHITE J., HESTON L., BIRD T.D., SCHELLENBERG G.D.;
RT   "Missense mutations in the chromosome 14 familial Alzheimer's disease
RT   presenilin 1 gene.";
RL   Hum. Mutat. 11:216-221(1998).
RN   [18]
RP   VARIANT AD GLU-378.
RX   MEDLINE; 99211215.
RA   BESANCON R., LORENZI A., CRUTS M., RADAWIEC S., STURTZ F.,
RA   BROUSSOLLE E., CHAZOT G., VAN BROECKHOVEN C., CHAMBA G.,
RA   VANDENBERGHE A.;
RT   "Missense mutation in exon 11 (codon 378) of the presenilin-1 gene in
RT   a French family with early-onset Alzheimer's disease and transmission
RT   study by mismatch enhanced allele specific amplification.";
RL   Hum. Mutat. 11:481-481(1998).
RN   [19]
RP   VARIANTS AD LEU-169 AND GLN-436.
RX   MEDLINE; 99047368.
RA   TADDEI K., KWOK J.B., KRIL J.J., HALLIDAY G.M., CREASEY H.,
RA   HALLUPP M., FISHER C., BROOKS W.S., CHUNG C., ANDREWS C.,
RA   MASTERS C.L., SCHOFIELD P.R., MARTINS R.N.;
RT   "Two novel presenilin-1 mutations (Ser169Leu and Pro436Gln) associated
RT   with very early onset Alzheimer's disease.";
RL   NeuroReport 9:3335-3339(1998).
RN   [20]
RP   VARIANT AD PRO-169.
RX   MEDLINE; 99148656.
RA   EZQUERRA M., CARNERO C., BLESA R., GELPI J.L., BALLESTA F., OLIVA R.;
RT   "A presenilin 1 mutation (Ser169Pro) associated with early-onset AD
RT   and myoclonic seizures.";
RL   Neurology 52:566-570(1999).

  Query Match             62.8%;  Score 1467;  DB 1;  Length 467;
  Best Local Similarity   65.7%;  Pred. No. 3.5e-94;
  Matches  306;  Conservative   39;  Mismatches   79;  Indels   42;  Gaps    8;

Qy      24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP-----GRPPG---- 75
           | | | |   :  |  ||   :   ||| :  : :|  ||    | |     ||| |
Db       3 ELPAPLSYFQNAQMSEDNHLSNTVRSQNDNRERQEHNDRR-SLGHPEPLSNGRPQGNSRQ 61

Qy      75 -------LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTED 127
                  :|||||||||||||||||||||:||||||||| ||| |:||||||||||
Db      62 VVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTED 121

Qy     128 TPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIY 187
           | :|||| |:|:||  |||||||||| |||||||||| || |||:|||:||| |::||
Db     122 TETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIY 181

Qy     188 LGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYL 247
           |||| ||||||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||
Db     182 LGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYL 241

Qy     248 PEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKL 307
           |||:||:||  |||||||||||||||||||||||||| :|||||||| ||| | ||:
Db     242 PEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEG 301

Qy     308 DPSSQGAL--QLPYDPE-MEEDSYDSFGE---PSYPEVFEPPLTGYPG------------ 350
           || :|  :     |: |  | :| |:  |     : | :|     : |
Db     302 DPEAQRRVSKNSKYNAESTERESQDTVAENDDGGFSEEWEAQRDSHLGPHRSTPESRAAV 361

Qy     350 EEL-------EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLT 402
           :||       |: ||||||||||||||||||||||:|| |||||||:||||||||||||
Db     362 QELSSSILAGEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLT 421

Qy     403 LLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           |||||:||||||||||||||:|||:|| ||:|||| || || ||
Db     422 LLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467


RESULT   8
PSN1_RAT
ID   PSN1_RAT       STANDARD;      PRT;   468 AA.
AC   P97887; P97529;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   PRESENILIN 1 (PS-1) (S182 PROTEIN).
GN   PSEN1 OR PSNL1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=BRAIN;
RX   MEDLINE; 97199371.
RA   TANIGUCHI T., HASHIMOTO T., TANIGUCHI R., SHIMADA K., KAWAMATA T.,
RA   YASUDA M., NAKAI M., TERASHIMA A., KOIZUMI T., MAEDA K., TANAKA C.;
RT   "Cloning of the cDNA encoding rat presenilin-1.";
RL   Gene 186:73-75(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=BRAIN;
RX   MEDLINE; 96255262.
RA   TAKAHASHI H., MURAYAMA M., TAKASHIMA A., MERCKEN M., NAKAZATO Y.,
RA   NOGUCHI K., IMAHORI K.;
RT   "Molecular cloning and expression of the rat homologue of
RT   presenilin-1.";
RL   Neurosci. Lett. 206:113-116(1996).
CC   -!- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE
CC       EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE.
CC       MAY FUNCTION IN THE CYTOPLASMIC PARTITIONING OF PROTEINS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D82578; BAA11575.1; -.
DR   EMBL; D82363; BAA11564.1; -.
DR   PFAM; PF01080; Presenilin; 1.
KW   Transmembrane; Glycoprotein.
FT   TRANSMEM     83    103       POTENTIAL.
FT   TRANSMEM    133    153       POTENTIAL.
FT   TRANSMEM    161    181       POTENTIAL.
FT   TRANSMEM    195    215       POTENTIAL.
FT   TRANSMEM    221    241       POTENTIAL.
```

```
PSN1_MICMU
ID   PSN1_MICMU     STANDARD;      PRT;   467 AA.
AC   P79802;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   PRESENILIN 1 (PS-1).
GN   PSEN1 OR PSNL1 OR PS1.
OS   Microcebus murinus (Lesser mouse lemur).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Strepsirhini; Cheirogaleidae; Microcebus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 97079199.
RA   CALENDA A., MESTRE-FRANCES N., CZECH C., PRADIER L., BONS N.,
RA   BELLIS M.;
RT   "Molecular cloning, sequencing, and brain expression of the
RT   presenilin 1 gene in Microcebus murinus.";
RL   Biochem. Biophys. Res. Commun. 228:430-439(1996).
CC   -!- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE
CC       EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE.
CC       MAY FUNCTION IN THE CYTOPLASMIC PARTITIONING OF PROTEINS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS; I-467 (SHOWN HERE) AND I-
CC       463; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: FOUND PREDOMINANTLY IN NEURONS OF THE
CC       DIFFERENT CORTICAL LAYERS AND HIPPOCAMPUS BUT ALSO IN SUBCORTICAL
CC       STRUCTURES.
CC   -!- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z71333; CAA95930.1; -.
DR   PFAM; PF01080; Presenilin; 1.
KW   Transmembrane; Alternative splicing; Glycoprotein.
FT   TRANSMEM     83    103       POTENTIAL.
FT   TRANSMEM    133    153       POTENTIAL.
FT   TRANSMEM    161    181       POTENTIAL.
FT   TRANSMEM    191    211       POTENTIAL.
FT   TRANSMEM    221    241       POTENTIAL.
FT   TRANSMEM    244    264       POTENTIAL.
FT   TRANSMEM    281    301       POTENTIAL.
FT   TRANSMEM    408    428       POTENTIAL.
FT   TRANSMEM    433    453       POTENTIAL.
FT   CARBOHYD    279    279       POTENTIAL.
FT   CARBOHYD    405    405       POTENTIAL.
FT   VARSPLIC     26     29       MISSING (IN ISOFORM I-463).
SQ   SEQUENCE   467 AA;  52384 MW;  A841A0B7 CRC32;


  Query Match             61.9%;  Score 1447;  DB 1;  Length 467;
  Best Local Similarity   64.4%;  Pred. No. 8.4e-93;
  Matches  300;  Conservative   41;  Mismatches   83;  Indels   42;  Gaps    7;

Qy      24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP-----GRPPG---- 75
           | | | |  :  |  ||   :   ||| :  : ::  ||    | |     ||| |
Db       3 ELPAPLSYFQNAQMSEDNHLSNTVRSQNDNREQQDHGDRRRL-GNPEPLSNGRPQGNSGP 61

Qy      75 -------LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTED 127
                  :||||||||||||||||||||||:||||||||| ||| |:||||||||||
Db      62 VVERDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTED 121

Qy     128 TPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIY 187
           | :|||| |:|||| ||||||||||| |||||||||||| || |||:|||:||| |::||
Db     122 TETVGQRALHSVLNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLIISSLLLLFFFSFIY 181

Qy     188 LGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYL 247
           |||| |||||:|| |: | :|||| |||: ||||||| ||||||||||||||||||||
Db     182 LGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYL 241

Qy     248 PEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKL 307
           |||:||:||  |||||||||||||||||||||||||||| :|||||||| ||| | ||:
Db     242 PEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWLVNMAEG 301

Qy     308 DPSSQGAL--QLPYD------------PEMEEDSYDSFGEPSYPEVFEP----PLTGYPG 349
           || :|  :     |:             || ::  :    |        |     ::
Db     302 DPEAQRRVSKNTKYNAQGTEREAQASVPENDDGGFSEEWEAQRDSQLGPHRSTSVSRAAV 361

Qy     350 EEL-------EEEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLT 402
           :|:       |: |||||||||||:|||||||||:|| |||||||:|||||||||||
Db     362 QEISSSIPASEDPEERGVKLGLGDFVFYSVLVGKASATASGDWNTTIACFVAILIGLCLT 421

Qy     403 LLLLAVFKKALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           |||||:|||||||||||||||:|||:|| ||:|||| || || ||
Db     422 LLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467


RESULT  11
PSN_DROME
ID   PSN_DROME      STANDARD;      PRT;   541 AA.
AC   O02194; O02395; O76802;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   PRESENILIN HOMOLOG (DPS) (DMPS).
GN   PS.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE FROM N.A. (LONG ISOFORM).
RX   MEDLINE; 97285868.
RA   BOULIANNE G.L., LIVNE-BAR I., HUMPHREYS J.M., LIANG Y., LIN C.,
RA   ROGAEV E., ST GEORGE-HYSLOP P.;
RT   "Cloning and characterization of the Drosophila presenilin
RT   homologue.";
RL   NeuroReport 8:1025-1029(1997).
RN   [2]
RP   SEQUENCE FROM N.A. (SHORT ISOFORM).
RC   STRAIN=CANTON-S;
RA   HONG C.S., KOO E.H.;
RT   "Isolation and characterization of Drosophila presenilin homolog.";
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
RA   YE Y., FORTINI M.E.;
RT   "Characterization of Drosophila presenilin and its colocalization
RT   with Notch during development.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U77934; AAB61139.1; -.
DR   EMBL; U78084; AAB53369.1; -.
DR   EMBL; AF084184; AAC33129.1; -.
DR   EMBL; AF084184; AAC33128.1; -.
DR   FLYBASE; FBgn0019947; PS.
DR   PFAM; PF01080; Presenilin; 1.
KW   Transmembrane; Glycoprotein; Alternative splicing.
FT   TRANSMEM    107    127       POTENTIAL.
```

```
GN   SEL-12.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC   Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE; 96032531.
RA   LEVITAN D., GREENWALD I.;
RT   "Facilitation of lin-12-mediated signalling by sel-12, a
RT   Caenorhabditis elegans S182 Alzheimer's disease gene.";
RL   Nature 377:351-354(1995).
RN   [2]
RP   REVISIONS TO 84-85.
RA   LEVITAN D.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAY FACILITATE LIN-12 MEDIATED RECEPTION OF
CC       INTERCELLULAR SIGNALS. IT MIGHT BE DIRECTLY INVOLVED IN LIN-12
CC       MEDIATED RECEPTION, FUNCTIONING AS A CO-RECEPTOR OR AS A
CC       DOWNSTREAM EFFECTOR THAT IS ACTIVATED UPON LIN-12 ACTIVATION.
CC       ALTERNATIVELY IT MAY BE INVOLVED IN A MORE GENERAL CELLULAR
CC       PROCESS SUCH AS RECEPTOR LOCALIZATION OR RECYCLING AND HENCE
CC       INFLUENCE LIN-12 ACTIVITY INDIRECTLY.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U35660; AAA85511.1; -.
DR   PFAM; PF01080; Presenilin; 1.
KW   Transmembrane.
FT   TRANSMEM     46     66       POTENTIAL.
FT   TRANSMEM    102    122       POTENTIAL.
FT   TRANSMEM    131    151       POTENTIAL.
FT   TRANSMEM    164    184       POTENTIAL.
FT   TRANSMEM    190    210       POTENTIAL.
FT   TRANSMEM    213    233       POTENTIAL.
FT   TRANSMEM    251    271       POTENTIAL.
FT   TRANSMEM    385    405       POTENTIAL.
FT   MUTAGEN      60     60       C->S: AR131; EGG-LAYING-DEFECTIVE.
SQ   SEQUENCE   461 AA;  51817 MW;  ABD50D85 CRC32;


  Query Match             42.1%;  Score 983;  DB 1;  Length 461;
  Best Local Similarity   49.1%;  Pred. No. 7.3e-61;
  Matches  222;  Conservative   60;  Mismatches   92;  Indels   78;  Gaps   12;

Qy      26 PTPRSCQEGRQGPEDGE-NTAQWRSQENEEDGEEDPDRYVCSGVPGRPPGLEEELTLKYG 84
           |: |  |||  |  | | :|    :     : :|| :             : ||  ||||
Db       2 PSTRRQQEG--GGADAETHTVYGTNLITNRNSQEDEN------------VVEEAELKYG 46

Qy      85 AKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQ-LIYTPFTEDTPSVGQRLLNSVLNTL 143
           | ||| |||||:||| :|| |: :: ||:: ||: |:|||| :| |: :: | |: | |
Db      47 ASHVIHLFVPVSLCMALVVFTMNTITFYSQNNGRHLLYTPFVRETDSIVEKGLMSLGNAL 106

Qy     144 IMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYNVAMDYPT 203
           :|: |:|:||: |:| ||| || |||||:|| :|||||| ||: ||||:::|:
Db     107 VMLCVVVLMTVLLIVFYKYKFYKLIHGWLIVSSFLLLFLFTTIYVQEVLKSFDVSPSALL 166

Qy     204 LLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVILGAISVYD 263
           :|  : |:| :||:|||||||| ||| ||| :|||||||||||||: | :|  |||:|
Db     167 VLFGLGNYGVLGMMCIHWKGPLRLQQFYLITMSALMALVFIKYLPEWTVWFVLFVISVWD 226

Qy     264 LVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMAKLDPSSQGALQLPYDPEM 323
           ||||| |||||| |||||||||||||||||||| :::   :         |::   ||
Db     227 LVAVLTPKGPLRYLVETAQERNEPIFPALIYSSGVIYPYVLVT-------AVENTTDPR- 279

Qy     324 EEDSYDSF------GE-------PSYPEVFEPP------------LTGYPGEELE----- 354
           |  | ||       ||       |  |:|  |             |  |  :|
Db     279 EPTSSDSNTSTAFPGEASCSSETPKRPKVKRIPQKVQIESNTTASTTQNSGVRVERELAA 338

Qy     354 ------------EEEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLC 400
                       ||||||||||||||||||||:|||::    |||||:||:||||||||
Db     339 ERPTVQDANFHRHEEEERGVKLGLGDFIFYSVLLGKASS--YFDWNTTIACYVAILIGLC 396

Qy     401 LTLLLLAVFKKALPALPISITFGLIFYFSTDN 432
           ||:||||||:|||||         | || |:
Db     397 FTLVLLAVFKRALPALQ--------FPFSPDS 420


RESULT  14
HOP1_CAEEL
ID   HOP1_CAEEL     STANDARD;      PRT;   358 AA.
AC   O02100;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-DEC-1999 (Rel. 39, Last annotation update)
DE   INTEGRAL MEMBRANE PROTEIN HOP-1.
GN   HOP-1 OR C18E3.8.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC   Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE; 98004548.
RA   LI X., GREENWALD I.;
RT   "HOP-1, a Caenorhabditis elegans presenilin, appears to be
RT   functionally redundant with SEL-12 presenilin and to facilitate LIN-12
RT   and GLP-1 signaling.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:12204-12209(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   CONNELL M., MAGGI L.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAY FACILITATE LIN-12 AND GLP-1 MEDIATED RECEPTION OF
CC       INTERCELLULAR SIGNALS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF021905; AAB84394.1; -.
DR   EMBL; AF000265; AAB52948.1; -.
DR   WORMPEP; C18E3.8; CE08317.
DR   PFAM; PF01080; Presenilin; 1.
KW   Transmembrane.
FT   TRANSMEM     13     33       POTENTIAL.
FT   TRANSMEM     58     78       POTENTIAL.
FT   TRANSMEM     87    107       POTENTIAL.
FT   TRANSMEM    116    136       POTENTIAL.
FT   TRANSMEM    149    169       POTENTIAL.
FT   TRANSMEM    171    191       POTENTIAL.
FT   TRANSMEM    298    318       POTENTIAL.
FT   TRANSMEM    322    342       POTENTIAL.
SQ   SEQUENCE   358 AA;  39864 MW;  FC6305ED CRC32;


  Query Match             22.5%;  Score 524.5;  DB 1;  Length 358;
  Best Local Similarity   30.5%;  Pred. No. 2.1e-29;
  Matches  117;  Conservative   81;  Mismatches  136;  Indels   49;  Gaps    8;

Qy      83 YGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPF----TEDTPSVGQRLLNS 138
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 18, 2000, 22:05:40 ; Search time 49.26 Seconds
(without alignments)
630.566 Million cell updates/sec

Title: US-08-509-359B-137
Perfect score: 2336
Sequence: 1 MLTFMASDSEEEVCDERTSL.....STDNLVRPFMDTLASHQLYI 448

Scoring table: BLOSUM62

Searched: 225878 seqs, 69334122 residues

Database : SPTREMBL_12:*

Word size : 0

Number of hits that pass the threshold : 225878
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2262.5	96.9	449	6	Q9XT96	Q9xt96 bos taurus
2	1438.5	61.6	478	6	Q9XT97	Q9xt97 bos taurus
3	1407.5	60.3	456	13	Q9W6T7	Q9w6t7 brachydanio
4	1355	58.0	384	13	O73869	O73869 cyprinus ca
5	576.5	24.7	272	5	O96340	O96340 drosophila
6	440	18.8	184	4	O95465	O95465 homo sapien
7	113.5	4.9	406	5	Q19737	Q19737 caenorhabdi
8	112.5	4.8	4578	13	O42181	O42181 fugu rubrip
9	110.5	4.7	320	8	Q34086	Q34086 coccyzus er
10	110.5	4.7	381	8	Q35425	Q35425 phascolosor
11	109	4.7	1840	11	O70611	O70611 rattus norv
12	108.5	4.6	380	8	Q9ZZC9	Q9zzc9 upupa epops
13	107.5	4.6	748	2	Q9Z577	Q9z577 streptomyce
14	107	4.6	381	8	Q33723	Q33723 antechinus
15	106.5	4.6	382	8	Q34340	Q34340 didelphis m
16	106.5	4.6	318	11	O35294	O35294 rattus norv
17	106.5	4.6	303	11	P97829	P97829 rattus norv
18	106	4.5	652	5	Q93346	Q93346 caenorhabdi
19	106	4.5	382	8	Q34279	Q34279 didelphis a
20	106	4.5	382	8	Q34677	Q34677 glironia ve
21	106	4.5	1717	13	Q90519	Q90519 fugu rubrip
22	105	4.5	444	2	Q9X2N3	Q9x2n3 arthrobacte
23	105	4.5	379	8	Q34428	Q34428 echimys did
24	104	4.5	382	8	Q35561	Q35561 philander o
25	104	4.5	379	8	Q36096	Q36096 trinomys pa
26	104	4.5	379	8	Q34430	Q34430 echimys did
27	104	4.5	826	10	O80739	O80739 arabidopsis
28	103	4.4	266	8	Q34429	Q34429 echimys did
29	103	4.4	266	8	Q37099	Q37099 echimys did
30	103	4.4	380	8	Q9XNM7	Q9xnm7 microtus lo
31	102.5	4.4	382	8	Q34973	Q34973 micoureus d
32	102.5	4.4	382	8	Q35075	Q35075 marmosa rub
33	102.5	4.4	382	8	Q35076	Q35076 micoureus r
34	102.5	4.4	408	8	Q9XP22	Q9xp22 marmota cau
35	102	4.4	444	2	O85080	O85080 arthrobacte
36	102	4.4	266	8	Q34425	Q34425 echimys did
37	102	4.4	266	8	Q34427	Q34427 echimys did
38	102	4.4	379	8	Q34424	Q34424 echimys did
39	102	4.4	379	8	Q33940	Q33940 akodon uric
40	102	4.4	381	8	Q9XP75	Q9xp75 sminthopsis
41	101.5	4.3	1681	11	Q62467	Q62467 mus musculu
42	101	4.3	270	3	O74710	O74710 emericella
43	101	4.3	381	8	Q9XP88	Q9xp88 sminthopsis
44	101	4.3	380	8	Q9XNL9	Q9xnl9 ondatra zib
45	100.5	4.3	382	8	Q35518	Q35518 philander m

ALIGNMENTS

RESULT 1
Q9XT96
ID Q9XT96 PRELIMINARY; PRT; 449 AA.
AC Q9XT96;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE PRESENILIN 2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae;
OC Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA SAHARA N., SHIRASAWA T., MORI H.;
RT "Molecular cloning of bovine presenilin 2 gene.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF038937; AAD39024.1; -.
SQ SEQUENCE 449 AA; 50301 MW; A3DA878F CRC32;

Query Match 96.9%; Score 2262.5; DB 6; Length 449;
Best Local Similarity 97.1%; Pred. No. 4.7e-162;
Matches 436; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

```
Qy       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQEGRQGPEDGENTAQWRSQENEED-GEED 59
           ||||||||||||||||||||||||||||||:|||| |||||: |||||||:|||  |||
Db       1 MLTFMASDSEEEVCDERTSLMSAESPTPRSCQDGRQGLEDGESAAQWRSQESEEDHEEED 60

Qy      60 PDRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQL 119
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 PDRYVCSGVPGRPPGLEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQL 120

Qy     120 IYTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLML 179
           |||||:|||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db     121 IYTPFSEDTPSVGQRLLNSVLNTLIMISVIVTMTIFLVVLYKYRCYKFIHGWLIMSSLML 180

Qy     180 LFLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALM 239
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     181 LFLFTYIYLGEVLKTYNVAMDYPTLFLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALM 240

Qy     240 ALVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMV 299
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ALVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMV 300

Qy     300 WTVGMAKLDPSSQGALQLPYDPEMEEDSYDSFGEPSYPEVFEPPLTGYPGEELEEEEERG 359
           |||||||||||||||||||||||||||||||||||:|||||| ||||||||||||||||
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RA   ARCHER S., HIRANO J., DISS J.K., FRASER S.P., DJAMGOZ M.B.A.;
RL   NeuroReport 0:0-0(0).
DR   EMBL; Y17128; CAA76641.1; -.
DR   PFAM; PF01080; Presenilin; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   384 AA;  43276 MW;  21A78D17 CRC32;


  Query Match             58.0%;  Score 1355;  DB 13;  Length 384;
  Best Local Similarity   69.9%;  Pred. No. 3.9e-94;
  Matches  270;  Conservative   33;  Mismatches   55;  Indels   28;  Gaps    5;

Qy      89 IMLFVPVTLCMIVVVATIKSVRFYTEKNG-QLIYTPFTEDTPSVGQRLLNSVLNTLIMIS 147
           ||||:||||||:|||||||||| |||:|:| ||||||| ||| :|||| |||:|| ||||
Db       1 IMLFIPVTLCMVVVVATIKSVSFYTQKDGQQLIYTPFREDTETVGQRALNSMLNAAIMIS 60

Qy     148 VIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEVLKTYNVAMDYPTLLLT 207
           ||||||: |||||||||| | |||  |:|:||| |::|||||| ||||||||| || :
Db      61 VIVVMTLVLVVLYKYRCYKVIQGWLFFSNLLLLFFFSFIYLGEVFKTYNVAMDYFTLAVI 120

Qy     208 VWNFGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVILGAISVYDLVAV 267
           :|||| |||:|||||||| ||||||||||||||||||||||:||:|| |||||||:||
Db     121 IWNFGVVGMICIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLILAAISVYDLLAV 180

Qy     268 LCPKGPLRMLVETAQERNEPIFPALIYSSAMVWTVGMA-KLDPSSQGALQLP-------- 319
           ||||||||:|||||||||||||||||| |||   ||   :| :  :  :|
Db     181 LCPKGPLRILVETAQERNEPIFPALIYSSTMVWLFNMADSAEPRNNSSHPVPQQENQDAV 240

Qy     319 ----------YDPEMEEDSYDSFG-----EPSYPEVFEPPLTGYPGEELEEEEERGVKL 362
                    : |          |     | |  |: | |     |   :|::|||||||
Db     241 APTAQPEDDGGFTPAWVNQQQHQLGPMQSTEDSRREIQELPSARPP--PVEDDEERGVKL 298

Qy     363 GLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITF 422
           |||||||||:|||||:|| |||||||:||||||||||||||||:|||||||||||||
Db     299 GLGDFIFYSMLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITF 358

Qy     423 GLIFYFSTDNLVRPFMDTLASHQLYI 448
           ||:|||:|||||||||| || || ||
Db     359 GLVFYFATDNLVRPFMDQLAVHQFYI 384


RESULT   5
O96340
ID   O96340      PRELIMINARY;      PRT;   272 AA.
AC   O96340;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   PRESENILIN (FRAGMENT).
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CANTON S;
RX   MEDLINE; 98331525.
RA   MARFANY G., DEL-FAVERO J., VALERO R., DE JONGHE C., WOODROW S.,
RA   HENDRIKS L., VAN BROECKHOVEN C., GONZALEZ-DUARTE R.;
RT   "Identification of a Drosophila presenilin homologue: evidence of
RT   alternatively spliced forms.";
RL   J. Neurogenet. 12:41-54(1998).
DR   EMBL; AF017025; AAD01611.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   272 AA;  29456 MW;  606B9A5C CRC32;


  Query Match             24.7%;  Score 576.5;  DB 5;  Length 272;
  Best Local Similarity   47.8%;  Pred. No. 5.7e-36;
  Matches  133;  Conservative   29;  Mismatches   41;  Indels   75;  Gaps    8;

Qy     240 ALVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVETAQERNEPIFPALIYSSAMV 299
           ||||||||||:|| :| |||::||:||| |:||||:|||||||||| |||||||| :|
Db       1 ALVFIKYLPEWTAWAVLAAISIWDLIAVLSPRGPLRILVETAQERNEQIFPALIYSSTVV 60

Qy     300 WTV----------GMAKLDPSS----------QGALQLPYDPEMEEDSYDSFGEPSYPE- 339
           : :          |   |||          | :|     ||    |    |  |:|
Db      61 YALVNTVTPQQSQATASSSPSSSNSTTTTRATQNSLA---SPEAAAASGQRTGN-SHPRQ 116

Qy     339 ---------VFEPPLTGYPGE---ELEE------------------------------ 355
                    |    |:  |    | |
Db     117 NQRDDGSVLATEAEAAGFTQEWSANLSERVARRQIEVQSTQSGNAQRSNEYRTVTAPDQN 176

Qy     355 ----EEERGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFK 410
                :||||:|||||||||||||||::   ||| ||:|||||||||||||||:::
Db     177 HPDGQEERGIKLGLGDFIFYSVLVGKASS--YGDWTTTIACFVAILIGLCLTLLLLAIWR 234

Qy     411 KALPALPISITFGLIFYFSTDNLVRPFMDTLASHQLYI 448
           |||||||||||||||| |:|  :|:|||: |:: |::|
Db     235 KALPALPISITFGLIFCFATSAVVKPFMEDLSAKQVFI 272


RESULT   6
O95465
ID   O95465      PRELIMINARY;      PRT;   184 AA.
AC   O95465;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   MINILIN.
GN   PSN1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   POWELL C.S., GEGG M.E., PALMER M.S.;
RT   "Human presenilin 1 gene encodes an alternative protein-minilin.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ008005; CAA07825.1; -.
SQ   SEQUENCE   184 AA;  21073 MW;  5C6FBAEE CRC32;


  Query Match             18.8%;  Score 440;  DB 4;  Length 184;
  Best Local Similarity   56.0%;  Pred. No. 6.2e-26;
  Matches  102;  Conservative   15;  Mismatches   43;  Indels   22;  Gaps    4;

Qy      24 ESPTPRSCQEGRQGPEDGENTAQWRSQENEEDGEEDPDRYVCSGVP-----GRPPG---- 75
           | | | |  :  |  ||   :   ||| :  : :|  ||    | |    ||| |
Db       3 ELPAPLSYFQNAQMSEDNHLSNTVRSQNDNRERQEHNDRR-SLGHPEPLSNGRPQGNSRQ 61

Qy      75 -------LEEELTLKYGAKHVIMLFVPVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTED 127
                  :||||||||||||||||||||:||||||| ||| |:||||||||||
Db      62 VVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYTPFTED 121

Qy     128 TPSVGQRLLNSVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIY 187
           | :|||| |:|:||  |||||||||| |||||||||||     : |    ||    :::
Db     122 TETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYK-----VSMRHRSLLSTLFFLW 176

Qy     188 LG 189
           ||
Db     177 LG 178


RESULT   7
Q19737
ID   Q19737      PRELIMINARY;      PRT;   406 AA.
AC   Q19737; Q22692;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   F22E10.5 PROTEIN.
GN   F22E10.5.
OS   Caenorhabditis elegans.
```

```
Qy      97 LCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFL 156
           ||:|  : |           | |:   :| ||            ||   ||
Db       7 LCLITQIVT-----------GLLLAMHYTADT------------TLAFSSV-------- 35

Qy     157 VVLYKYRCYKFIHGWLI----MSSLMLLFLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFG 212
             : |    :||||     :   : |:  |:::|     |   ::  |     || |
Db      35 ----AHTCRNVQYGWLIRNLHANGASMFFICIYLHIGRGF-YYGSYLNKET-----WNTG 84

Qy     213 AVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVILGAISVYDLVAVLCPKG 272
                          :|::  |||  |: |:  |      ||  : :| : :   |
Db      85 ------------------VILLLTLMATAFVGYVLPWGQMSFWGATVITNLFSAIPYIG 125

Qy     273 PLRMLVETAQ---ERNEP----------IFPALIYSSAMVWTVGMAKLDPSSQGALQ--- 317
             : ||| |     : |          : | :|    :|    : :   ::   ||
Db     126 --QTLVEWAWGGFSVDNPTLTRFFALHFLLPFMIAGLTIVHLTFLHETGSNNPLGLQSNC 183

Qy     317 --LPYDPEMEEDSYDSFG-----------EPSY---PEVFEP--PLTGYPGEELEEEEE 357
             : : |          |              |:    || | |  ||   |
Db     184 DKISFHPYFSLKDLVGFTITLLLLLTTLALFSPTLLGDPENFSPANPLVTTP--------- 235

Qy     358 RGVKLGLGDFIFYSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALP 417
             :|     |  |: |       |       ||   ::|: | |  ||    |:|:|: |
Db     235 -HIKPDWSFFFEYATLRSIPIKLGG-----VLALKESVLV-LSLAPLLKNSKKRAMPSRP 287

Qy     418 ISITFGLIFYFSTDNLVRPFMDTLASH 444
           :| |   :|:|  :||    :  | ||
Db     288 LSQT---LFWFLVENLF--ILKWLGSH 309


RESULT  10
Q35425
ID   Q35425      PRELIMINARY;      PRT;   381 AA.
AC   Q35425;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   CYTOCHROME B.
OS   Phascolosorex dorsalis.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Metatheria; Dasyuromorphia; Dasyuridae; Phascolosorex.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93096825.
RA   KRAJEWSKI C., DRISKELL A.C., BAVERSTOCK P.R., BRAUN M.J.;
RT   "Phylogenetic relationships of the thylacine (Mammalia: Thylacinidae)
RT   among dasyuroid marsupials: evidence from cytochrome b DNA
RT   sequences.";
RL   Proc. R. Soc. Lond., B, Biol. Sci. 250:19-27(1992).
CC   -!- CATALYTIC ACTIVITY: QH(2) + 2 FERRICYTOCHROME C = Q + 2
CC       FERROCYTOCHROME C.
CC   -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC       BOUND TO THE PROTEIN (BY SIMILARITY).
DR   EMBL; M99462; AAB61692.1; -.
DR   PFAM; PF00033; cytochrome_b_N; 1.
KW   Mitochondrion; Electron transport; Respiratory chain; Transmembrane;
KW   Heme.
SQ   SEQUENCE   381 AA;  42813 MW;  C6A2674F CRC32;


  Query Match             4.7%;  Score 110.5;  DB 8;  Length 381;
  Best Local Similarity   21.1%;  Pred. No. 0.7;
  Matches   81;  Conservative   54;  Mismatches  115;  Indels  133;  Gaps   18;

Qy      97 LCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFL 156
           :|:|: : |           | :   :| ||                    :| |
Db      39 MCLIIQILT-----------GLFLAMHYTSDT----------------------LTAFT 64

Qy     157 VVLYKYRCYKFIHGWLIM------SSLMLLFLFTYIYLGEVLKTYNVAMDYPTLLLTVWN 210
             |   : |    :|||:       :|:  : || :|  |    :|        |    ||
Db      65 SV--AHICRDVNYGWLLRNLHANGASMFFMCLFLHIGRGIYYGSY--------LYKETWN 114

Qy     211 FGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVILGAISVYDLVAVLCP 270
           |                    :|:: |:||  |: |:  |      ||  : :|:: :
Db     115 IG------------------VILLLAVMATAFVGYVLPWGQMSFWGATVITNLLSAIPY 155

Qy     271 KGPLRMLVE-------------TAQERNEPIFPALIYSSAMVWTVGMAKL---DPS--SQ 312
           ||   | |              |      | | :: : |:|  : : :    :||  :
Db     156 IGP--TLAEWVWGGYAVDKATLTRFFAFHFILPFIVTALAIVHLLFLHETGSNNPSGINP 213

Qy     313 GALQLPYDPE-----------------MEEDSYDSFGEPSYPEVFEP--PLTGYPGEEL 352
           | ::|: |                   :   | || |:   |: | |  ||   |  :
Db     214 NADKIPFHPYYTIKDALGFMLLLSVLLLLTLFSPDSLGD---PDNFSPANPLNTPPHIKP 270

Qy     353 EEEEERGVKLGLGDFIF-YSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAVFKK 411
           |              |:| |::|       |       ||   :||| | : ||  |  ::
Db     271 E-----------WYFLFAYAILRSIPNKLGG-----VLALLASILILLIIPLLHTA-NQR 313

Qy     412 ALPALPISITFGLIFYFSTDNLV 434
           ::   ||| |   :|:  | ||:
Db     314 SMMFRPISQT---LFWILTANLI 333


RESULT  11
O70611
ID   O70611      PRELIMINARY;      PRT;  1840 AA.
AC   O70611;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   RAT SKELETAL MUSCLE TYPE 1 VOLTAGE-GATED SODIUM CHANNEL
DE   (RSKM1) VARIANT (RSKM1).
GN   SCN4A.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COPENHAGEN; TISSUE=PROSTATE;
RA   DISS J.K., STEWART D., BLACK J.A., FRASER S.P., DIBB-HAJJ S.,
RA   WAXMAN S.G., ARCHER S., DJAMGOZ M.B.A.;
RL   FEBS Lett. 0:0-0(0).
DR   EMBL; Y17153; CAA76659.1; -.
DR   PFAM; PF00520; ion_trans; 4.
DR   PFAM; PF00612; IQ; 1.
DR   PRINTS; PR00170; NACHANNEL.
KW   Ionic channel.
SQ   SEQUENCE   1840 AA;  208823 MW;  1948B0C2 CRC32;


  Query Match             4.7%;  Score 109;  DB 11;  Length 1840;
  Best Local Similarity   20.3%;  Pred. No. 5.4;
  Matches   98;  Conservative   63;  Mismatches  130;  Indels  192;  Gaps   28;

Qy       5 MASDSEEEVCDERTSLM------SAESPTPRSCQEGRQGPE-DGENTAQWRSQENEE--- 55
           :| | |:|  :|   ::         |    :: |    | | ||: |       |::
Db     452 LAEDQEKE--EEFQQMLEKYKKHQEELEKAKAAQALESGEEADGDPT------HNKDCNG 503

Qy      55 ----DGEEDPDRYVCSGVPGRPPGLEE----------------ELTLKYGA-------KH 87
               ||: | |  ||        :||                    | :          ||
Db     504 SLDASGEKGPPRPSCSADSAISDAMEELEEAHQKCPPWWYKCAHKVLIWNCCAPWVKFKH 563

Qy      88 VIMLFV-------PVTLCMIVVVATIKSVRFYTEKNGQLIYTPFTE---DTPSVGQRLLN 137
           :| | |         :|:|  ||:      ::  |           | ||    :  |||  :
Db     564 IIYLIVMDPFVDLGITIC-IVLNTLFMAMEHY----------PMTEHFDNVLSVGNLVFT 612

Qy     138 SVLNTLIMISVIVVMTIFLVVLYKYRCYKFIHGWLIMSSLMLLFLFTYIYLGEV------ 192
           :    ::: :| :         |:|    |  || |  | ::       : |  |
Db     613 GIFTAEMVLKLIAMDP------YEY----FQQGWNIFDSFIVTLSLVELGLANVQGLSVL 662

Qy     192 -----LKTYNVAMDYPTLLLTVWNFG-AVGMVCIHWKGPLVLQQAYLIMISALMAL-VFI 244
                |: : :|  :||| : :   | :|| :     | | |  | :: | |:: : :|
Db     663 RSFRLLRVFKLAKSWPTLNMLIKIIGNSVGAL-----GNLTLVLAIIVFIFAVVGMQLFG 717

Qy     245 K-------------YLPEW----------------SAWV--------ILG---AISVYD 263
```

```
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   CYTOCHROME B.
OS   Antechinus naso (Antechinus habbema).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Metatheria; Dasyuromorphia; Dasyuridae; Antechinus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   KRAJEWSKI C., PAINTER J., DRISKELL A.C., BUCKLEY L., WESTERMAN M.;
RL   Sci. New Guinea 19:157-166(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   KRAJEWSKI C., PAINTER J., BUCKLEY L., WESTERMAN M.;
RL   J. Mammal. Evol. 2:25-35(1994).
RN   [3]
RP   SEQUENCE OF 1-287 FROM N.A.
RA   KRAJEWSKI C., BUCKLEY L., WOOLLEY P.A., WESTERMAN M.;
RL   J. Mammal. Evol. 3:81-91(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   ARMSTRONG L.A., KRAJEWSKI C., WESTERMAN M.;
RL   J. Mammal. 0:0-0(1998).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   KRAJEWSKI C.;
RL   Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RA   KRAJEWSKI C.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U07576; AAC03634.1; -.
DR   PFAM; PF00032; cytochrome_b_C; 1.
DR   PFAM; PF00033; cytochrome_b_N; 1.
KW   Mitochondrion.
SQ   SEQUENCE   381 AA;  42738 MW;  9C6998A0 CRC32;


  Query Match             4.6%;  Score 107;  DB 8;  Length 381;
  Best Local Similarity   21.3%;  Pred. No. 1.3;
  Matches   76;  Conservative   53;  Mismatches  111;  Indels  116;  Gaps   17;

Qy    138 SVLNTLIMISVIVVMTIFLVVLY-----------KYRCYKFIHGWLIM------SSLMLL 180
          |:|   ::| ::  | :|| : |            : |    |||||       :|:  :
Db     35 SLLGACLIIQIL--MGLFLAMHYTSDTLTAFSSVAHICRDVNHGWLIHNLHVNGASMFFM 92

Qy    181 FLFTYIYLGEVLKTYNVAMDYPTLLLTVWNFGAVGMVCIHWKGPLVLQQAYLIMISALMA 240
            || ::  |    :|        |   || |                  :|::  :||
Db     93 CLFLHMGRGIYYGSY--------LYKETWNIG------------------VILLLTVMA 125

Qy    241 LVFIKYLPEWSAWVILGAISVYDLVAVLCPKGPLRMLVE------------TAQERNEP 287
            |: |:  |      ||  : :|:: :   |    | |             |
Db    126 TAFVGYVLPWGQMSFWGATVITNLLSAIPYIG--TTLAEWIWGGFAVDKATLTRFFAFHF 183

Qy    288 IFPALIYSSAMVWTVGM--------AKLDPSSQGALQLPYDPE----------------- 323
          | | :: : |:|  : :         : ::| |    ::|: |
Db    184 ILPFIVVALAIVHLLFLHETGSNNPSGINPDSD---KIPFHPYYTIKDALGMTLLLLALL 240

Qy    323 -MEEDSYDSFGEPSYPEVFEP--PLTGYPGEELEEEEERGVKLGLGDFIF-YSVLVGKAA 378
           :   | || |:   |: | |  ||   |  : |             |:| |::|
Db    241 LLALFSPDSLGD---PDNFSPANPLNTPPHIKPE----------WYFLFAYAILRSIPN 286

Qy    379 ATGSGDWNTTLACFVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLV 434
            |       ||   :||| | : ||  |  ::::   ||| |   :|:  | ||:
Db    287 KLGG-----VLALLASILILLIIPLLHTA-NQRSMMFRPISQT---LFWILTANLI 333


RESULT  15
Q34340
ID   Q34340      PRELIMINARY;      PRT;   382 AA.
AC   Q34340;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   CYTOCHROME B LIGHT STRAND.
OS   Didelphis marsupialis.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Metatheria; Didelphimorphia; Didelphidae; Didelphis.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   PATTON J.L., REIS MARIA S.F., DA SILVA N.F.;
RL   J. Mammal. Evol. 3:3-29(1996).
CC   -!- CATALYTIC ACTIVITY: QH(2) + 2 FERRICYTOCHROME C = Q + 2
CC       FERROCYTOCHROME C.
CC   -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC       BOUND TO THE PROTEIN (BY SIMILARITY).
DR   EMBL; U34665; AAA99746.1; -.
DR   PFAM; PF00032; cytochrome_b_C; 1.
DR   PFAM; PF00033; cytochrome_b_N; 1.
KW   Mitochondrion; Electron transport; Respiratory chain; Transmembrane;
KW   Heme.
SQ   SEQUENCE   382 AA;  43139 MW;  907FBCA3 CRC32;


  Query Match             4.6%;  Score 106.5;  DB 8;  Length 382;
  Best Local Similarity   20.5%;  Pred. No. 1.4;
  Matches   79;  Conservative   53;  Mismatches  115;  Indels  139;  Gaps   18;

Qy     97 LCMIVVVATIKSVRFYTEKNGQLIYTPFTEDTPSVGQRLLNSVLNTLIMISVIVVMTIFL 156
          :|:|: : |           |  :   :| ||                        :| |
Db     39 MCLIIQILT-----------GLFLAMHYTSDT----------------------LTAFS 64

Qy    157 VVLYKYRCYKFIHGWLIM------SSLMLLFLFTYIYLGEVLKTYNVAMDYPTLLLTVWN 210
           |   : |    :||||       :|:  : || ::  |    :|        |    ||
Db     65 SV--AHICRDVNYGWLIRNIHANGASMFFMCLFLHVGRGIYYGSY--------LYKETWN 114

Qy    211 FGAVGMVCIHWKGPLVLQQAYLIMISALMALVFIKYLPEWSAWVILGAISVYDLVAVLCP 270
           |                    :|::  :||  |: |:  |       ||  : :|:: :
Db    115 IG------------------VILLLTVMATAFVGYVLPWGQMSFWGATVITNLLSAIPY 155

Qy    271 KGPLRMLVE------------TAQERNEPIFPALIYSSAMVWTVGMAK--------LDP 309
           |    |||              |       | | :| :  :|  : : :        |||
Db    156 IG--NTLVEWIWGGFSVDKATLTRFFAFHFILPFIILAMVVVHLLFLHETGSNNPTGLDP 213

Qy    310 SSQGALQLPYDP-----------------EMEEDSYDSFGEPSYPEVFEP--PLTGYPG 349
          :|    ::|: |                   :   | |  |:   |: | |  ||   |
Db    214 NSD---KIPFHPYYTIKDILGLFLMIIILLSLAMFSPDLLGD---PDNFTPANPLNTPPH 267

Qy    350 EELEEEEERGVKLGLGDFIF-YSVLVGKAAATGSGDWNTTLACFVAILIGLCLTLLLLAV 408
           : |             |:| |::|       |       ||   :||| | : ||  :
Db    268 IKPE----------WYFLFAYAILRSIPNKLGG-----VLALLASILILLIMPLLHTST 311

Qy    409 FKKALPALPISITFGLIFYFSTDNLV 434
           ::::   ||| |   :|:  | ||:
Db    312 -QRSMMFRPISQT---LFWMLTANLI 333


Search completed: March 18, 2000, 22:07:43
Job time: 123 sec
```

```
=> e

E6            1     PRESENILIN-2 PS2S (MOUSE PS-2SHORT ISOFORM)/CN
E7            1     PRESENILIN-ASSOCIATED PROTEIN 1 (HPAP-1) (HUMAN INCYTE
CLONE
                     1353337)/CN
E8            1     PRESENILINASE/CN
E9            1     PRESEP-AGRI/CN
E10           1     PRESER ACE/CN
E11           1     PRESERCAR/CN
E12           1     PRESERIN M 72/CN
E13           1     PRESERIN T 72/CN
E14           1     PRESERT/CN
E15           1     PRESERVAC WETPROOF/CN
E16           1     PRESERVAL/CN
E17           1     PRESERVAL B/CN

=> s e4-6

              1 "PRESENILIN-2 (HUMAN ISOFORM)"/CN
              1 "PRESENILIN-2 PS2CCAS (MOUSE)"/CN
              1 "PRESENILIN-2 PS2S (MOUSE PS-2SHORT ISOFORM)"/CN
L1            3 ("PRESENILIN-2 (HUMAN ISOFORM)"/CN OR "PRESENILIN-2 PS2CCAS
                (MOUSE)"/CN OR "PRESENILIN-2 PS2S (MOUSE PS-2SHORT
ISOFORM)"/CN)

=> d 1-3 ide can


L1   ANSWER 1 OF 3  REGISTRY  COPYRIGHT 2000 ACS
RN   251358-30-2  REGISTRY
CN   Presenilin-2 (human isoform) (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   5: PN: WO9960122 SEQID: 5 claimed protein
FS   PROTEIN SEQUENCE
MF   Unspecified
CI   MAN
SR   CA
LC   STN Files:   CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
               1 REFERENCES IN FILE CA (1967 TO DATE)
               1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

REFERENCE   1:  132:11416

L1   ANSWER 2 OF 3  REGISTRY  COPYRIGHT 2000 ACS
RN   200445-64-3  REGISTRY
CN   330-448-presenilin-2 (Mus musculus isoform PS2Ccas) (9CI)  (CA INDEX
NAME)
OTHER NAMES:
CN   Presenilin-2 PS2Ccas (mouse)
FS   PROTEIN SEQUENCE
MF   Unspecified
CI   MAN
SR   CA
LC   STN Files:   CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
```